

日本国特許庁
PATENT OFFICE
JAPANESE GOVERNMENT

JC523 U.S. PTO
09/048026
03/26/98

別紙添付の書類に記載されている事項は下記の出願書類に記載されて
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
with this Office.

出願年月日
Date of Application:

1997年 9月 8日

出願番号
Application Number:

平成 9年特許願第242247号

出願人
Applicant(s):

富士通株式会社

CERTIFIED COPY OF
PRIORITY DOCUMENT

1997年12月 5日

特許庁長官
Commissioner,
Patent Office

荒井寿光

出証番号 出証特平09-3099178

【書類名】 特許願

【整理番号】 9701389

【提出日】 平成 9年 9月 8日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 17/00

【発明の名称】 関連文書表示装置

【請求項の数】 10

【発明者】

 【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

 【氏名】 内野 寛治

【発明者】

 【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

 【氏名】 津田 宏

【発明者】

 【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

 【氏名】 松井 くにお

【特許出願人】

 【識別番号】 000005223

 【氏名又は名称】 富士通株式会社

【代理人】

 【識別番号】 100074099

 【郵便番号】 102

 【住所又は居所】 東京都千代田区二番町8番地20 二番町ビル3F

 【弁理士】

 【氏名又は名称】 大菅 義之

 【電話番号】 03-3238-0031

【選任した代理人】

【識別番号】 100067987

【郵便番号】 222

【住所又は居所】 神奈川県横浜市港北区太尾町1418-305（大倉
山二番館）

【弁理士】

【氏名又は名称】 久木元 彰

【電話番号】 045-545-9280

【手数料の表示】

【予納台帳番号】 012542

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9705047

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 関連文書表示装置

【特許請求の範囲】

【請求項1】 参照関係を有する文書からなる文書群を表示する関連文書表示装置であって、

文書集合を構成する複数の文書間の参照関係を解析することによって、前記文書集合を構成する複数の文書を参照関係を有する1つ以上の文書群に分類し、その分類に関する情報を文書群情報として抽出する文書群解析手段と、

前記文書群情報を参照することにより、前記文書群毎に、その文書群を構成する文書に含まれるキーワードを抽出する文書群キーワード抽出手段と、

前記文書群毎に、それに関連するタイトルと、該文書群に対して前記文書群キーワード抽出手段によって抽出されたキーワードとを対応付けて表示する文書群キーワード表示手段と、

を含むことを特徴とする関連文書表示装置。

【請求項2】 請求項1に記載の装置であって、

前記文書群キーワード表示手段は、ユーザが指定した検索指定に該当する文書が含まれる文書群又は該検索指定に該当するキーワードを強調表示する、

ことを特徴とする関連文書表示装置。

【請求項3】 参照関係を有する文書からなる文書群を表示する関連文書表示装置であって、

文書集合を構成する複数の文書間の参照関係を解析することによって、前記文書集合を構成する複数の文書を参照関係を有する1つ以上の文書群に分類し、その分類に関する情報を文書群情報として抽出する文書群解析手段と、

前記文書集合を構成する複数の文書のそれぞれから、その文書の属性を示す文書属性情報を抽出する文書属性解析手段と、

前記文書群情報及び前記文書属性情報を参照することによって、前記文書群毎に、その文書群の参照関係を、その文書群を構成する各文書に対応する各文書属性情報又はその省略された情報が各ノードとして表示されるツリー構造によって表示する文書群構造表示手段と、

を含むことを特徴とする関連文書表示装置。

【請求項4】 請求項3に記載の装置であって、

前記文書群構造表示手段は、前記文書群情報及び前記文書属性情報を参照することによって、前記文書群毎に、その文書群の参照関係を、その文書群を構成する各文書に対応する各省略された作者名と各文書量情報とが組み合せられて各ノードとして表示されるツリー構造によって表示する、

ことを特徴とする関連文書表示装置。

【請求項5】 請求項3に記載の装置であって、

前記文書群構造表示手段は、前記文書群毎に、その文書群に含まれる文書から抽出される複数のトピックを、その文書群に対応して表示される前記ツリー構造を構成する前記各ノードと関連付けて、更に表示する、

ことを特徴とする関連文書表示装置。

【請求項6】 請求項5に記載の装置であって、

前記文書群構造表示手段は、前記文書群毎に、前記各トピックとそれに関連する前記各ノードとを、同一の色で表示する、

ことを特徴とする関連文書表示装置。

【請求項7】 請求項3に記載の装置であって、

前記文書群構造表示手段は、ユーザが指定した検索指定に該当する文書に対応するノードを強調表示する、

ことを特徴とする関連文書表示装置。

【請求項8】 参照関係を有する文書からなる文書群を表示する関連文書表示装置であって、

文書集合を構成する複数の文書間の参照関係を解析することによって、前記文書集合を構成する複数の文書を参照関係を有する1つ以上の文書群に分類し、その分類に関する情報を文書群情報として抽出する文書群解析手段と、

該分類された各文書群を、それぞれを構成する文書から抽出されるトピックに基づいて更に細分類し、その細分類に関する情報をトピック分類情報として抽出するトピック解析手段と、

前記文書群情報及び前記トピック分類情報を参照することにより、前記文書群

毎及びそれを細分類した前記トピック毎に、そのトピックに関連する文書に含まれるキーワードを抽出するトピックキーワード抽出手段と、

前記文書群毎及びそれを細分類した前記トピック毎に、それに関連するタイトルと、該トピックに対して前記トピックキーワード抽出手段によって抽出されたキーワードとを対応付けて表示するトピックキーワード表示手段と、

を含むことを特徴とする関連文書表示装置。

【請求項9】 コンピュータにより使用されたときにそれによって読み出されるプログラムを記録した記録媒体であって、

文書集合を構成する複数の文書間の参照関係を解析することによって、前記文書集合を構成する複数の文書を参照関係を有する1つ以上の文書群に分類し、その分類に関する情報を文書群情報として抽出する機能と、

前記文書群情報を参照することにより、前記文書群毎に、その文書群を構成する文書に含まれるキーワードを抽出する機能と、

前記文書群毎に、それに関連するタイトルと、該文書群に対して前記文書群キーワード抽出手段によって抽出されたキーワードとを対応付けて表示する機能と

を前記コンピュータに行わせるためのプログラムを記録したコンピュータ読出し可能記録媒体。

【請求項10】 コンピュータにより使用されたときにそれによって読み出されるプログラムを記録した記録媒体であって、

文書集合を構成する複数の文書間の参照関係を解析することによって、前記文書集合を構成する複数の文書を参照関係を有する1つ以上の文書群に分類し、その分類に関する情報を文書群情報として抽出する機能と、

前記文書集合を構成する複数の文書のそれぞれから、その文書の属性を示す文書属性情報を抽出する機能と、

前記文書群情報及び前記文書属性情報を参照することによって、前記文書群毎に、その文書群の参照関係を、その文書群を構成する各文書に対応する各文書属性情報又はその省略された情報が各ノードとして表示されるツリー構造によって表示する機能と、

を前記コンピュータに行わせるためのプログラムを記録したコンピュータ読み可能記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、参照関係にある文書群を整理し、その文書群をいろいろな観点から表示することによって、ユーザによる必要な情報へのアクセスを支援する技術に関する。

【0002】

【従来の技術及び発明が解決しようとする課題】

パソコン通信又はコンピュータネットワーク上で運用される電子会議室や電子ニュース等において、順次蓄積される文書集合の中から必要な文書をより迅速かつ簡単に見つけ出したいという要請が、従来からある。

【0003】

このような要請に対して、文書集合中の各文書のタイトルを作成日順に並び替え、その結果得られるタイトルリストをユーザに掲示するという従来技術が知られている。

【0004】

また、文書集合を互いに参照関係のある文書から構成される文書群に分類し、文書群中の各文書のタイトルをインデントして表示することにより、各文書の参照関係を掲示する従来技術や、文書群中の各文書の番号をツリーで表示することにより各文書の参照関係を掲示する従来技術も知られている。

【0005】

更に、文書集合の中から、特定のキーワードを含む文書を全文検索し、その検索結果を羅列的に掲示する従来技術も知られている。

しかし、これらの従来技術におけるような限定された情報表示のみでは、以下に示される問題点を解決することができなかった。

1. 雑多な文書集合の中から必要な文書又は必要な文書が含まれているであろう文書群を見つけ出すためには、掲示される文書のタイトルに頼るしかない

い。タイトルは、必ずしも文書内容を正確に表わしているとは限らないため、正確な検索が困難である。

2. タイトルのインデント表示や文書番号のツリー表示だけでは、文書群全体の構造を把握すること、及び文書群における話題の推移を把握することが困難である。
3. 種々の観点から必要な文書にアクセスすることができない。
4. 検索結果が多数件ある場合に、更に絞り込み検索を実行するか検索結果のリストを1件1件チェックしなければ、必要な文書にアクセスすることができない。

一方、複数の特定の文書からキーワードを抽出し、共通のキーワードを含む各文書に対し自動的に他の文書へのリンクを設定する技術が、知られている。この従来技術は、特許の公知例や研究論文等の特定文書中で互いの文献を相互参照することを可能にすることによって、関連する複数の文書を効率的に読み広げることが可能にする。

【0006】

しかし、このような従来技術は、関連する文書の参照を容易にすることを目的としており、電子会議室や電子ニュース等の文書集合からの、必要な文書又は必要な文書が含まれているであろう文書群の検索の支援に、適用することはできなかった。

【0007】

本発明の課題は、大量の文書が含まれる文書集合からの、必要な文書又は必要な文書が含まれているであろう文書群の検索を、支援することにある。

【0008】

【課題を解決するための手段】

本発明は、参照関係を有する文書からなる文書群を表示する関連文書表示装置を前提とする。

【0009】

本発明の第1の態様は、以下の構成を有する。

まず、文書群解析手段（構造解析部402）は、文書集合を構成する複数の文

書間の参照関係を解析することによって、文書集合を構成する複数の文書を参照関係を有する1つ以上の文書群（スレッド）に分類し、その分類に関する情報を文書群情報（スレッドインデックス305）として抽出する。

【0010】

文書群キーワード抽出手段（内容解析部403）は、文書群情報を参照することにより、文書群毎に、その文書群を構成する文書に含まれるキーワードを抽出する。

【0011】

文書群キーワード表示手段（表示装置306）は、文書群毎に、それに関連するタイトルと、その文書群に対して文書群キーワード抽出手段によって抽出されたキーワードとを対応付けて表示する。

【0012】

上述した本発明の第1の態様の構成により、キーワードを頼りにして、雑多な文書集合の中から必要な文書が含まれているであろう文書群を容易に見つけ出すことが可能となる。

【0013】

本発明の第2の態様は、上述の本発明の第1の態様の構成において、文書群キーワード表示手段が、ユーザが指定した検索指定に該当する文書が含まれる文書群又はその検索指定に該当するキーワードを強調表示する構成を有する。

【0014】

上述した本発明の第2の態様の構成により、検索指定に該当する文書群を一目で把握することが可能となる。

本発明の第3の態様は、以下の構成を有する。

【0015】

文書群解析手段（構造解析部402）は、文書集合を構成する複数の文書間の参照関係を解析することによって、文書集合を構成する複数の文書を参照関係を有する1つ以上の文書群に分類し、その分類に関する情報を文書群情報として抽出する。

【0016】

文書属性解析手段（書式解析部401）は、文書集合を構成する複数の文書のそれぞれから、その文書の属性を示す文書属性情報（メタインデックス304）を抽出する。

【0017】

文書群構造表示手段（表示装置306）は、文書群情報及び文書属性情報を参照することによって、文書群毎に、その文書群の参照関係を、その文書群を構成する各文書に対応する各文書属性情報又はその省略された情報が各ノードとして表示されるツリー構造によって表示する。より具体的には、文書群構造表示手段は、文書群毎に、その文書群の参照関係を、その文書群を構成する各文書に対応する各省略された作者名と各文書量情報（各行数）とが組み合せられて各ノードとして表示されるツリー構造によって表示する。

【0018】

上述した本発明の第3の態様の構成により、文書群全体の構造を容易に把握することが可能となる。

本発明の第4の態様は、上述の本発明の第3の態様の構成において、文書群構造表示手段が、文書群毎に、その文書群に含まれる文書から抽出される複数のトピック（メイントピック、サブトピック）を、その文書群に対応して表示されるツリー構造を構成する各ノードと関連付けて、更に表示する構成を有する。より具体的には、文書群構造表示手段は、文書群毎に、各トピックとそれに関連する各ノードとを、同一の色で表示する。

【0019】

上述した本発明の第4の態様の構成により、文書群全体の話題の推移を容易に把握することが可能となる。

本発明の第5の態様は、前述の本発明の第3の態様の構成において、文書群構造表示手段が、ユーザが指定した検索指定に該当する文書に対応するノードを強調表示する構成を有する。

【0020】

上述した本発明の第5の態様の構成により、スレッドの構造を把握しつつ、検索を実行することが可能となる。

本発明の第6の態様は、以下の構成を有する。

【0021】

文書群解析手段（構造解析部402）は、文書集合を構成する複数の文書間の参照関係を解析することによって、文書集合を構成する複数の文書を参照関係を有する1つ以上の文書群に分類し、その分類に関する情報を文書群情報として抽出する。

【0022】

トピック解析手段（構造解析部402）は、その分類された各文書群を、それぞれを構成する文書から抽出されるトピックに基づいて更に細分類し、その細分類に関する情報をトピック分類情報（サブトピックインデックス）として抽出する。

【0023】

トピックキーワード抽出手段（内容解析部403）は、文書群情報及びトピック分類情報を参照することにより、文書群毎及びそれを細分類したトピック毎に、そのトピックに関連する文書に含まれるキーワードを抽出する。

【0024】

トピックキーワード表示手段（表示装置306）は、文書群毎及びそれを細分類したトピック毎に、それに関連するタイトルと、そのトピックに対してトピックキーワード抽出手段により抽出されたキーワードとを対応付けて表示する。

【0025】

上述した本発明の第6の態様の構成により、キーワードによるより精密なトピックの絞込みを行うことが可能となる。

なお、本発明は、コンピュータにより使用されたときに、上述の本発明の各構成によって実現される機能と同様の機能をコンピュータに行わせるためのコンピュータ読出し可能記録媒体として構成することもできる。

【0026】

【発明の実施の形態】

以下、図面を参照しながら本発明の実施の形態について詳細に説明する。

全体構成

図1は、本発明の実施の形態が対象とする、文書集合及び文書群の例を示す図である。

【0027】

この例では、コンピュータネットワーク上に、ユーザが話題別に議論を行うフォーラムと呼ばれる仮想的な公開討論会場が設けられており、各フォーラムは、会議室と呼ばれる、更に細分化された話題を扱う複数の仮想的な会場に分類されている。ユーザがこの会議室に、発言文書を投稿（アップロード）することによって、議論が進行する。フォーラム及び会議室とも、サーバコンピュータ上のストレージエリアとして構成され、それらには、上述した分類基準に従って、文書が蓄積される。また、各会議室において、互いに参照関係を有する複数の文書からなる文書群がスレッドを構成する。

【0028】

ユーザが投稿する文書は、例えば、図2に示されるデータ構造を有しており、その文書の番号を示す文書番号、日付、タイトル、その文書が参照する文書の番号である参照文書番号、作者名（発言者名）等の文書の属性フィールドが記載されるヘッダ部と、文書の本体が記載される内容部とから構成されている。

【0029】

本発明の実施の形態では、以下のような表示形態が可能となる。

1. キーワードビュー：会議室を構成する文書集合において、その文書集合中の各スレッド毎に、そのスレッドを構成する文書群からキーワードが抽出され、それらのキーワードが、それらの文書数及びそれらが含まれるスレッドのタイトルと合わせて、図25に示される表示形態で表示される。

【0030】

キーワードビューによって、ユーザは、キーワードを頼りにして、雑多な文書集合の中から必要な文書が含まれているであろうスレッド（文書群）を容易に見つけ出すことが可能となる。

2. スレッドビュー：文書の参照関係、タイトル、作者名、及び行数が一目にわかる図26に示される表示形態で、各スレッドを構成する文書群が表示される。

【0031】

スレッドビューにより、スレッド全体の構造を把握し話題の推移を容易に把握することが可能となる。

3. 発言者ビュー：各文書のタイトルが、発言者（作者）毎に分類され、かつ発言者が発言の多い順にソートされ、同一発言者内では日付順で、図27に示される表示形態で、表示される。

【0032】

発言者及び発言日付という観点から、文書集合（会議室）内の文書を参照することが可能となる。

4. 各ビューへの検索結果の反映：ユーザが指定した検索キーワードに関連する文書が、図32又は図33に示される表示形態で、キーワードビュー、スレッドビュー等の表示中で強調表示される。

【0033】

この結果、より正確な文書の把握が可能となる。

5. 各ビューの切替え機能：上述のキーワードビュー、スレッドビュー、及び発言者ビューが任意に切替え可能とされることにより、種々の観点から必要な文書にアクセス可能となる。

以上のような表示形態を可能とする本発明の実施の形態について、詳細に説明する。

【0034】

図3及び図4は、本発明の実施の形態のシステム構成図である。

フォーラム／会議室内の文書群は、所定のサーバコンピュータ上の文書群データベース301として、蓄積される。

【0035】

文書群解析装置302は、文書群データベース301内の各会議室に対応する文書集合毎に、それに含まれる文書群の解析を行う。

集計装置303は、文書群解析装置302による解析結果に基づいて、メタインデックス304、スレッドインデックス305、及び索引ファイル404を生成する。

【0036】

表示装置306は、メタインデックス304とスレッドインデックス305を用いて、キーワードビュー、スレッドビュー、又は発言者ビューの何れかの表示形態で、文書群を表示する。

【0037】

また、文字列検索装置405は、ユーザによる検索語の指定に基づいて、索引ファイル404を用いながら文書群データベース301内の文書集合を構成する各文書に対して検索を実行する。表示装置306は、その検索結果を、キーワードビュー又はスレッドビューに反映させて表示する。

【0038】

表示装置306は、メタインデックス304とスレッドインデックス305を用いて、キーワードビュー、スレッドビュー、又は発言者ビューの何れかの表示形態で、文書群を表示する。

【0039】

文書群解析装置302は、書式解析部401、構造解析部402、及び内容解析部403とから構成される。

書式解析部401は、文書群データベース301内の文書集合を構成する図2のデータ構造を有する各文書のヘッダ部から、文書番号、タイトル、作者名、日付、及び参照文書番号を抽出し、また、各文書の内容部の行数を算出し、それらを、集計装置303を経由して、図5に示されるデータ構造を有するメタインデックス304に登録する。

【0040】

構造解析部402は、書式解析部401が各文書から抽出した文書番号と参照文書番号に基づいて、各文書をスレッドを単位とする文書群に分類し、集計装置303を経由して、スレッド毎に、それを構成する文書の参照関係のリストであるスレッドインデックス305を作成する。

【0041】

図6は、スレッドインデックス305のデータ構造を示す図である。

各スレッド毎に、ルート文書番号と、文書数と、スレッドの構造を示すリスト

とが登録される。リストは、

(親文書番号 子文書番号/サブツリー 子文書番号/サブツリー....)

という記述形式によって記述され、”子文書番号/サブツリー”の部分には、更に再帰的(リカーシブ)に、子リストを記述することができる。

【0042】

図6に例示される2つのスレッドの各リストにより表現される各参照関係は、同図の表の右側に示される如くである。

また、構造解析部402は、解析したスレッドを構成する文書群を、更にタイトルを同一とするサブ文書群に分類し、各サブ文書群に色番号を付与し、各サブ文書群に含まれる文書に対応する図5のデータ構造を有するメタインデックス304内のエントリに、その文書が属するサブ文書群に付与された色番号を登録する。

【0043】

内容解析部403は、構造解析部402によって分類されたスレッド毎に、そのスレッドを構成する文書群を1つの結合文書ファイルにまとめ、その結合文書からキーワードを抽出する。日本語文書からキーワードを抽出する技術としては、種々の公知技術を採用することができる。この場合に、キーワード抽出の精度を向上させるために、ノイズとなる文字がパターンマッチングによって除去される。また、例えば、上位所定個数のキーワードのみが抽出される。

【0044】

内容解析部403によって抽出された各スレッドのキーワードは、集計装置303を経由して、そのスレッドのルート文書に対応する図5のデータ構造を有するメタインデックス304のエントリに、登録される。

【0045】

また、内容解析部403は、スレッド毎に抽出したキーワードから、そのキーワードに含まれる索引語を抽出し、集計装置303を経由して、図7に示されるデータ構造を有する索引ファイル404を生成する。

【0046】

この索引ファイル404は、前述したように、文字列検索装置405によって

参照される。

文書群解析装置 302 の詳細説明

図 8 は、図 4 の文書群解析装置 302 内の書式解析部 401 及び構造解析部 402 が実現する制御動作を示す動作フローチャートである。

【0047】

まず、書式解析部 401 は、文書群データベース 301 から新たに登録された新規文書ファイルから文書データを 1 行ずつ読み込みながら、その文書ファイルのヘッダ部（図 2 参照）から、文書番号、タイトル、作者名、日付、及び参照文書番号を抽出する（ステップ 801→802→803→801 のループ）。

【0048】

書式解析部 401 は、ヘッダ部の抽出を終了すると、集計装置 303 を経由して、図 5 のデータ構造を有するメタインデックス 304 において新規エントリを生成し、そのエントリに、抽出した文書番号、タイトル、作者名、日付、及び参照文書番号を登録する（ステップ 802→804）。

【0049】

次に、書式解析部 401 は、上記新規文書ファイル内のヘッダ部以降の内容部（図 2 参照）から文書データを 1 行ずつ読み込みながら、文書末尾（E O F : エンドオブファイル）が検出されるまで、内容部の行数をカウントする（ステップ 805→806→807→805 のループ）。

【0050】

書式解析部 401 は、文書末尾を検出すると、集計装置 303 を経由して、それまでにカウントした内容部の行数を、図 5 のデータ構造を有するメタインデックス 304 の、現在処理中の新規文書の文書番号に対応するエントリに、登録する（ステップ 806→808）。

【0051】

なお、上述の行数のカウント処理において、他の文書から引用している行（例えば”>”で始まる行）については、行数のカウントには算入しないことによって、その文書が実質的に発言している行数をカウントするように構成されてもよい。

【0052】

続いて、構造解析部402に制御が移り、構造解析部402は、まず、現在処理中の新規文書の文書番号を、集計装置303を経由して、図6のデータ構造を有するスレッドインデックス305に登録する（ステップ809）。

【0053】

図9は、上記ステップ809の登録動作を示す動作フローチャートである。

まず、構造解析部402は、現在処理中の新規文書が、或るスレッドのルート文書であるか否かを判定する（ステップ901）。具体的には、構造解析部402は、図8のステップ801～803のループにおいて、現在処理中の新規文書から参照文書番号が検出されなかった場合に、その文書はルート文書であると判定する。

【0054】

構造解析部402は、現在処理中の新規文書が或るスレッドのルート文書であると判定した場合には、集計装置303を経由して、スレッドインデックス305において新規エントリを生成し、そのエントリに現在処理中の新規文書の文書番号をルート文書番号として登録する（ステップ901→902）。

【0055】

構造解析部402は、ステップ902の処理の後、上記エントリの文書数を1に初期設定し（ステップ906）、図8のステップ809の処理を終了する。

一方、構造解析部402は、現在処理中の新規文書が或るスレッドのルート文書ではないと判定した場合には、現在処理中の新規文書のヘッダ部から抽出されている参照文書番号が含まれるスレッドインデックス305中のエントリに、その参照文書番号を親文書番号とするリストが存在するか否かを判定する（ステップ901→903）。

【0056】

構造解析部402は、上述のエントリに、現在処理中の新規文書のヘッダ部から抽出されている参照文書番号を親文書番号とするリストが存在すると判定した場合には、現在処理中の新規文書のヘッダ部から抽出されている文書番号を、そのリストの子文書番号として登録する（ステップ903→905）。

【0057】

一方、構造解析部402は、上述のエントリに、現在処理中の新規文書のヘッダ部から抽出されている参照文書番号を親文書番号とするリストが存在しないと判定した場合には、そのエントリに、その参照文書番号を親文書番号とするリストを生成した上で、現在処理中の新規文書のヘッダ部から抽出されている文書番号を、そのリストの子文書番号として登録する（ステップ903→904→905）。

【0058】

構造解析部402は、ステップ905の処理の後、上述のエントリの文書数を更新（プラス1）し（ステップ906）、図8のステップ809の処理を終了する。

【0059】

上記図9の動作フローチャートによって実現される制御動作の具体例につき、図10の説明図を用いて説明する。この図は、図6のスレッドインデックス305において、ルート文書番号が”001”であるスレッドのエントリのリストが生成される過程を示すものである。

【0060】

まず、文書番号”001”のルート文書が処理される時点で、図9のステップ901→902が実行されることにより、スレッドインデックス305において新規エントリが生成され、そのエントリに文書番号”001”がルート文書番号として登録され（図10の（1））、上記エントリの文書数が1に初期設定される。

【0061】

次に、文書番号”002”の文書が処理される時点で、図9のステップ901→903→904→905→906が実行されることにより、スレッドインデックス305内のルート文書番号”001”のエントリにおいて、文書番号”002”の文書から抽出された参照文書番号”001”を親文書番号とするリストが生成された後（図10の（2））、文書番号”002”がそのリストの子文書番号として登録され（図10の（3）の下線部）、上記エントリの文書数が2に更

新される。

【0062】

次に、文書番号”003”の文書が処理される時点で、図9のステップ901→903→904→905→906が実行されることにより、スレッドインデックス305内のルート文書番号”001”のエントリにおいて、文書番号”003”の文書から抽出された参照文書番号”002”を親文書番号とするリストが生成された後（図10の（4）の下線部）、文書番号”003”がそのリストの子文書番号として登録され（図10の（5）の下線部）、上記エントリの文書数が3に更新される。

【0063】

次に、文書番号”004”の文書が処理される時点で、図9のステップ901→903→905→906が実行されることにより、スレッドインデックス305内のルート文書番号”001”のエントリにおいて、文書番号”004”の文書から抽出された参照文書番号”001”を親文書番号とするリストの子文書番号として、文書番号”004”が登録され（図10の（6）の下線部）、上記エントリの文書数が4に更新される。

【0064】

次に、文書番号”005”の文書が処理される時点で、図9のステップ901→903→904→905→906が実行されることにより、スレッドインデックス305内のルート文書番号”001”のエントリにおいて、文書番号”005”の文書から抽出された参照文書番号”004”を親文書番号とするリストが生成された後（図10の（7）の下線部）、文書番号”005”がそのリストの子文書番号として登録され（図10の（8）の下線部）、上記エントリの文書数が5に更新される。

【0065】

最後に、文書番号”006”の文書が処理される時点で、図9のステップ901→903→904→905→906が実行されることにより、スレッドインデックス305内のルート文書番号”001”のエントリにおいて、文書番号”006”の文書から抽出された参照文書番号”005”を親文書番号とするリスト

が生成された後（図10の（9）の下線部）、文書番号”006”がそのリストの子文書番号として登録され（図10の（10）の下線部）、上記エントリの文書数が6に更新される。

【0066】

以上説明した図8のステップ809の処理の後、構造解析部402は、現在処理中の新規文書について色番号を決定し、その色番号を図5のデータ構造を有するメタインデックス304中の上記新規文書の文書番号に対応するエントリに登録する処理を実行する（図8のステップ810）。

【0067】

図11は上記ステップ810の登録動作を示す動作フローチャートである。なお、この登録動作では、図12に示されるデータ構造を有するカラーテーブルが使用される。このテーブルは、特には図示しない記憶装置に記憶される。

【0068】

まず、構造解析部402は、現在処理中の新規文書が、或るスレッドのルート文書であるか否かを判定する（ステップ1101）。具体的には、構造解析部402は、図8のステップ801～803のループにおいて、現在処理中の新規文書から参照文書番号が検出されなかった場合に、その文書はルート文書であると判定する。

【0069】

構造解析部402は、現在処理中の新規文書が或るスレッドのルート文書であると判定した場合は、そのルート文書の文書番号に対応するエントリを、図12のデータ構造を有するカラーテーブルに登録し、そのエントリに現在処理中の新規文書から抽出された文書番号及びタイトル（図8のステップ804参照）と、初期色番号を登録する（ステップ1101→1102→1103）。図12の例では、ルート文書番号”001”の色番号”#1”に対応するエントリが登録され、そのタイトルはメインピックとなり、また、そのエントリの文書番号フィールドには、当初はルート文書番号”001”のみが登録される。

【0070】

その後、構造解析部402は、図5のデータ構造を有するメタインデックス3

04 中の上記新規文書の文書番号に対応するエントリに、ステップ1103で登録した初期色番号を登録し（ステップ1103→1110）、図8のステップ810の処理を終了する。

【0071】

一方、構造解析部402は、現在処理中の新規文書が或るスレッドのルート文書ではないと判定した場合には、現在処理中の新規文書から抽出されたタイトル（図8のステップ804参照）が、“Re:”等の参照記号を含んでいるか否かを判定する（ステップ1101→1104）。

【0072】

構造解析部402は、現在処理中の新規文書から抽出されたタイトルが参照記号を含んでいると判定した場合はそのタイトルから参照記号を削除し（ステップ1104→1105）、現在処理中の新規文書から抽出されたタイトルが参照記号を含んでいないと判定した場合にはステップ1105は実行しない。

【0073】

その後、構造解析部402は、図12のデータ構造を有するカラーテーブル中の現在処理中の新規文書が属するスレッドに対応する何れかのエントリに、現在処理中の新規文書から抽出され参照記号を含まないタイトルと同じタイトルが登録されているか否かを判定する（ステップ1106）。現在処理中の新規文書が属するスレッドとそのルート文書番号は、図8のステップ809の処理において図6のデータ構造を有するスレッドインデックス305のエントリが決定される際に検出されるため、そのルート文書番号からカラーテーブル中のエントリが決定される。例えば、現在処理中の新規文書が文書番号“002”の文書である場合には、図12に示されるカラーテーブルにおいて、ルート文書番号“001”に属するエントリが検出される。

【0074】

構造解析部402は、図12のデータ構造を有するカラーテーブル中の現在処理中の新規文書が属するスレッドに対応する何れかのエントリに、現在処理中の新規文書から抽出され参照記号を含まないタイトルと同じタイトルが登録されていると判定した場合には、そのエントリの文書番号フィールドに、現在処理中の

新規文書の文書番号を登録する（ステップ1106→1107）。例えば、現在処理中の新規文書が文書番号”002”の文書である場合には、図12に示されるカラーテーブルにおいて、ルート文書番号”001”に属し色番号”#1”が登録されているエントリの文書番号フィールドに、文書番号”002”が登録される。

【0075】

その後、構造解析部402は、図5のデータ構造を有するメタインデックス304中の上記新規文書の文書番号に対応するエントリに、ステップ1107で登録が行われたカラーテーブル中のエントリに設定されている色番号を登録し（ステップ1107→1110）、図8のステップ810の処理を終了する。

【0076】

一方、構造解析部402は、図12のデータ構造を有するカラーテーブル中の現在処理中の新規文書が属するスレッドに対応する何れのエントリにも、現在処理中の新規文書から抽出され参照記号を含まないタイトルと同じタイトルが登録されてはいないと判定した場合は、カラーテーブルにおいて上記スレッドに対応する新たなエントリを作成し（ステップ1108）、その作成したエントリに、そのスレッド内で新たな色番号と、現在処理中の新規文書から抽出された文書番号及びタイトル（図8のステップ804参照）を登録する（ステップ1106→1108→1109）。例えば、現在処理中の新規文書が文書番号”003”の文書である場合には、図12のカラーテーブルにおいて、ルート文書番号”001”に属する新たなエントリが作成され、そのエントリに、色番号”#2”と、文書番号”003”の文書のタイトルと、文書番号”003”とが登録される。このタイトルは、ルート文書番号”001”のタイトルであるメインピックに対して、サブピック1となる。

【0077】

その後、構造解析部402は、図5のデータ構造を有するメタインデックス304中の上記新規文書の文書番号に対応するエントリに、ステップ1109でカラーテーブル中の新たなエントリに設定された新たな色番号を登録し（ステップ1109→1110）、図8のステップ810の処理を終了する。

【0078】

内容解析部403は、図6のデータ構造を有するスレッドインデックス305を参照することにより、前述したように、スレッド毎に、そのスレッドを構成する文書群を1つの結合文書ファイルにまとめ、その結合文書からキーワードを抽出する。この結果、抽出された各スレッドのキーワードは、そのスレッドのルート文書に対応する図5のデータ構造を有するメタインデックス304のエントリに、登録される。

表示装置306の詳細説明

表示装置306は、前述したように、図5のデータ構造を有するメタインデックス304と図6のデータ構造を有するスレッドインデックス305を用いて、キーワードビュー、スレッドビュー、又は発言者ビューの何れかの表示形態で、文書群を表示することができる。

【0079】

ここで例えば、図4のシステムが、ホームページの表示を制御するWebサーバに接続されるように構成されれば、ユーザは、パーソナルコンピュータ等の手元の端末上のWebブラウザアプリケーションから上記Webサーバに接続して特定のフォーラムの特定の会議室にログインした後に、所定の各GUI（グラフィックユーザインタフェース）ボタンをマウス装置等でクリックすることによって、キーワードビュー、スレッドビュー、又は発言者ビューを切り替えて表示させることができる。

【0080】

より具体的には、表示装置306は、Webサーバに対して例えばCGI（コモンゲートウェイインタフェース）アプリケーションとして機能し、Webサーバから引き渡されたユーザからのリクエストに応答して、キーワードビュー、スレッドビュー、又は発言者ビュー等の各ビューを表現するHTML（ハイパーテキストマークアップ言語）による文書データを生成し、それをWebサーバに引き渡す。そして、これらのHTML文書データをWebサーバがユーザにインターネット等のコンピュータネットワークを経由して返信することにより、ユーザの端末上のWebブラウザアプリケーションに、上記ビューが表示される。

【0081】

まず、表示装置306が実現するキーワードビューの表示動作について説明する。

前述したようにキーワードビューにおいては、スレッド毎に、そのスレッドを構成する文書群から抽出されているキーワードが、その文書群の文書数及びそのスレッドのタイトルと合わせて、図25に示される表示形態で表示される。

【0082】

図13は、表示装置306が実行するキーワードビューの表示動作を示す動作フローチャートである。

まず、表示装置306は、図5のデータ構造を有するメタインデックス304のファイルを読み込む（ステップ1301）。

【0083】

次に、表示装置306は、メタインデックス304のファイルから1エントリずつデータを読み込みながら、ルート文書が登録されているエントリを検索する（ステップ1301→1302→1301のループ）。各エントリがルート文書が登録されているエントリであるか否かは、各エントリの参照文書番号フィールドの値が無効なデータ値であるか否かによって判定することができる。

【0084】

表示装置306は、ルート文書が登録されているエントリを検出すると、そのルート文書番号を、そのルート文書番号に対応する文書群データベース301内のルート文書を表示するためのアプリケーションへの統一されたアドレス情報であるURL (Uniform Resource Locator) がHREF属性の値として指定されるアンカータグに変換する（ステップ1302→1303）。

【0085】

次に表示装置306は、図6のデータ構造を有するスレッドインデックス305において、上記ルート文書番号に対応するエントリを参照することにより、そのスレッドに含まれる文書数（子文書数）を取得する（ステップ1304）。

【0086】

そして、表示装置306は、図5のデータ構造を有するメタインデックス30

4において、上記ルート文書が登録されているエントリから、タイトル（メイントピック）と、キーワードとを抽出し、それらと、ステップ1303で変換されたアンカータグ形式のルート文書番号、及びステップ1304で取得した子文書数からなるデータ列を1テーブルレコードとして含むHTMLテーブル文書データを作成する（ステップ1305）。

【0087】

続いて、表示装置306は、メタインデックス304のファイルから文書末尾（EOF）を検出するまで、上記ステップ1301～1305の一連の処理を繰り返し実行することにより、各スレッド毎のHTMLテーブル文書データを作成する（ステップ1306→1301）。

【0088】

表示装置306は、メタインデックス304のファイルから文書末尾を検出すると（ステップ1306の判定がYES）、最終的に得られたHTMLテーブル文書データをWebサーバに引き渡して、キーワードビューの表示動作を終了する。この結果、ユーザの端末のWebブラウザアプリケーション上に、図25に例示されるようなテーブル形式で、キーワードビューが表示される。

【0089】

ユーザは、キーワードビュー上の各スレッド毎のキーワードを頼りにして、雑多な文書集合の中から必要な文書が含まれているであろうスレッドを容易に見つけ出すことが可能となる。

【0090】

また、ユーザは、ルート文書に対応するアンカーをマウス装置等でクリックすることによって、所望のスレッドのルート文書に即座にアクセスすることができる。

【0091】

上述のキーワードビューの表示動作において、子文書数に応じて、各スレッドのテーブルレコードを色分けして表示するように構成されてもよい。これによって、ユーザは、スレッド毎の発言数を一目で判別することができる。

【0092】

続いて、表示装置306が実現するスレッドビューの表示動作について説明する。

前述したように、スレッドビューにおいては、文書の参照関係、タイトル、作者名、及び行数が一目にわかる図26に示される表示形態で、各スレッドを構成する文書群が表示される。

【0093】

図26において、スレッドの参照関係及び話題の推移が色付きツリーによって表示される。各ツリーのノードは、各文書に対応し、その文書の作者名の先頭文字（2バイト）とその文書の行数を用いて、

作者名[行数]

の形式で表示される。また、各ノードの前後には、“*”、“+”、“=”、又は“.”等の記号が付される。これらの記号の意味は、下記の通りである。

“*” この記号が付される文書がルート文書である。

【0094】

“+” この記号が付される文書が参照している文書が他の文書によっても参照されている。

“=” この記号が付される文書を参照している文書が存在する。

【0095】

“.” この記号が付される文書を参照している文書が存在しない。

また、図26において、“Main Topic:”に続いてそのスレッドのルート文書のタイトルが表示され、“Sub Topic:”に続いてそのスレッド中に現れるルート文書のタイトル以外のタイトルが表示される。そして、各タイトルは色分けされ、各タイトルと同じタイトル（参照記号を除く）を有する文書に対応するノードは、そのタイトルの色と同じ色で表示される。

【0096】

これによって、ユーザは、スレッド全体の構造を把握しスレッド内の話題の推移を一目で把握することが可能となる。

更に、各ノードはアンカーとして表示される。これにより、ユーザは、各ノードをマウス装置等によってクリックすることにより、そのノードに対応する文書

に即座にアクセスすることができる。

【0097】

図14は、表示装置306が実行するスレッドビューの表示動作を示す動作フローチャートである。。

まず、表示装置306は、図6のデータ構造を有するスレッドインデックス305のファイルから、1つのスレッドに対応する1つのエントリ（1行）のリストと、そのスレッドに含まれる文書数を、読み込む（ステップ1401）。例えば、図6のデータ構造を有するスレッドインデックス305において、ルート文書番号”001”に対応するリストとして、

(001 (002 003) (004 (005 006)))

が読み込まれ、文書数として”6”が読み込まれる。

【0098】

次に、表示装置306は、読み込んだリストから、例えば図6の表の右側に示されるスレッドのツリー構造を復元する（ステップ1402）。このツリー構造を表現するために、表示装置306は、例えば図15に示されるような配列データを生成する。

【0099】

次に、表示装置306は、読み込んだリストの各ノードを構成する文書番号毎に、その文書番号に対応する図5のデータ構造を有するメタインデックス304のエントリを抽出し、そのエントリから、作者名、行数、色番号、及びタイトルを抽出する（ステップ1403）。これらの抽出されたデータは、上記各ノードに対応付けて記憶される。

【0100】

次に、表示装置306は、ステップ1401で読み込んだ文書数と、ステップ1403で抽出した各ノードの色番号とから、スレッドビューの先頭で表示される各タイトルの色を決定する（ステップ1404）。この動作は、各色番号に実際の色をマッピングする動作として実現される。

【0101】

次に、表示装置306は、スレッドに含まれるルート文書のタイトルとその他

のタイトルを、"Main Topic:" 及び"Sub Topic:" に続けて表示するためのHTML文書を作成する。この場合に、各タイトルは、前述した構造解析部402が管理する図12に示されるカラーテーブルの上記スレッドに属する各エントリから順次読み出され、同時に順次読み出される各色番号からステップ1404で決定された各色が算出され、その各色での表示が順次指定される。各色は、HTML文書の色指定命令（ タグ等）によって指定される。

【0102】

最後に、表示装置306は、ステップ1402で復元したスレッドのツリー構造を示す配列データを構成する左端のノードの文書番号から順に処理することにより、そのツリー構造を表示するためのHTML文書を作成する（ステップ1406）。この場合、前述したように、表示装置306は、ステップ1403で抽出した各ノードの作者名、行数、及び色番号に基づいて、ツリー構造の各ノードの文書番号を、そのノードに対応する文書の作者名の先頭文字（2バイト）とその文書の行数とからなる表示データ、

作者名 [行数]

に変換し、更に、その表示データをそのノードの色番号に対応する色で表示させるためのHTML文書データを生成する。色番号と実際の色との対応関係は、ステップ1404で決定された対応関係に従う。また、前述したように、表示装置306は、各ノードに対応する上記表示データの前後に、その接続関係に基づいて、" * "、" + "、" = "、又は" . " 等の記号を表示するためのHTML文書データを生成する。ここで、ツリー構造をそのままの形式で表示可能とするために、例えば、HTMLにおける制御用タグであるプリフォーマットタグ<PRE>が使用される。更に、上記ノード毎の表示データは、そのノードに対応する文書群データベース301内の文書データを表示するためのアプリケーションへのURLがHREF属性の値として指定されるアンカータグとして生成される。

【0103】

続いて、表示装置306は、スレッドインデックス305のファイルから文書末尾（EOF）を検出するまで、上記ステップ1401～1406の一連の処理

を繰り返し実行することにより、各スレッド毎のビューデータを作成する（ステップ1407→1401）。

【0104】

表示装置306は、スレッドインデックス305のファイルから文書末尾を検出すると（ステップ1407の判定がYES）、最終的に得られたHTMLテーブル文書データをWebサーバに引き渡して、スレッドビューの表示動作を終了する。この結果、ユーザの端末のWebブラウザアプリケーション上に、図26に例示されるような形式で、スレッドビューが表示される。

【0105】

次に、表示装置306が実現する発言者ビューの表示動作につき説明する。

前述したように、発言者ビューにおいては、各文書のタイトルが、発言者（作者）毎に分類され、かつ発言者が発言の多い順にソートされ、同一発言者内では日付順で、図27に示される表示形態で、表示される。

【0106】

図16は、表示装置306が実行する発言者ビューの表示動作を示す動作フローチャートである。。

表示装置306は、発言者ビューを実現するために、図17のデータ構造を有する作者配列データを使用する。そして、表示装置306は、発言者ビューの表示開始時に、この作者配列データを初期化する（ステップ1601）。

【0107】

次に、表示装置306は、図5のデータ構造を有するメタインデックス304のファイルから1つのエントリのデータを読み込む（ステップ1602）。

次に、表示装置306は、このエントリから抽出される作者名の作者が、作者配列データに含まれていない作者であるか否かを判定する（ステップ1603）

。

【0108】

表示装置306は、上記エントリから抽出される作者名の作者が、作者配列データに含まれていない作者である場合には、作者配列データに新しい作者項目を追加する（ステップ1603→1604）。表示装置306は、上記エントリか

ら抽出される作者名の作者が、作者配列データに含まれている作者である場合には、ステップ1604の処理は実行しない。

【0109】

次に、表示装置306は、作者配列データ中の該当する作者項目に、上記エントリから抽出される文書番号を登録する（ステップ1605）。

続いて、表示装置306は、メタインデックス304のファイルから文書末尾（EOF）を検出するまで、上記ステップ1602～1605の一連の処理を繰り返し実行することにより、メタインデックス304に登録されている全ての文書番号を、作者別に作者配列データに登録する。

【0110】

表示装置306は、メタインデックス304のファイルから文書末尾を検出すると（ステップ1606の判定がNO）、作者配列データ中の各作者項目を、それぞれの項目に登録されている文書番号の数、即ち各作者毎の発言文書数に基づいてソートする（ステップ1607）。

【0111】

続いて、表示装置306は、作者配列データ中の同一作者項目内で、文書番号を、それに対応するメタインデックス304中のエントリから抽出される日付に基づいてソートする（ステップ1608）。

【0112】

最後に、表示装置306は、上記ステップ1607及び1608でのソートの結果得られる作者配列データの各作者項目毎に、作者名と、その項目内の各文書番号に対応するメタインデックス304中のエントリから抽出される日付及びタイトルを表示するためのHTMLテーブル文書データを生成し、それをWebサーバに引き渡して、発言者ビューの表示動作を終了する。この結果、ユーザの端末のWebブラウザアプリケーション上に、図27に例示されるようなテーブル形式で、発言者ビューが表示される。

【0113】

ユーザは、発言者ビュー上で、発言者及び発言日付という観点から、文書集合（会議室）内の文書を参照することが可能となる。

また、或る発言者の発言を時間を追って参照したり、会議室内で多くの発言をするリーダー的な発言者を一目で確認することができる。

表示装置306の他の表示態様

次に、上記各ビューの表示動作以外に表示装置306が実現する各表示動作の態様について説明する。

【0114】

まず、表示装置306が実現する発言内容表示の動作につき説明する。

前述したように、ユーザは、キーワードビューにおけるそれぞれのスレッド上のアンカー又はスレッドビューにおける各ノード上のアンカーを、マウス装置等でクリックすることにより、各スレッドのルート文書又は各ノードに対応する文書等に、即座にアクセスすることができる。

【0115】

ユーザによってこれらの操作が実行された場合には、Webサーバから指示によって、表示装置306によって実行される図18に示される動作フローチャートの処理が例えばCGIとして起動される。この場合、この処理には、ユーザによって指定されたアンカータグに含まれる文書番号の情報が引き渡される。

【0116】

この結果まず、表示装置306は、上記文書番号の情報を読み込んだ後（ステップ1801）、ヘッダ部に上記読み込んだ文書番号と同じ文書番号を含んでいる文書ファイルを読み込むまで、文書群データベース301からの文書ファイルの読み込みを行う（ステップ1802→1803→1802のループ）。

【0117】

表示装置306は、ヘッダ部に上記読み込んだ文書番号と同じ文書番号を含んでいる文書ファイルを読み込むと（ステップ1803の判定がYES）、新しい文書のヘッダ部を読み込むまで、ステップ1804～1809のループにより、上記文書ファイルから1行ずつデータを読み込み、そのデータを1行分のHTML文書データに変換し、そのHTML文書データをWebサーバに出力する（ステップ1808）。

【0118】

この場合に、各行のデータが他の文書等へのURLを含んでいる場合には、表示装置306は、そのデータを上記URLがHREF属性の値として指定されるアンカータグに変換した上で出力する（ステップ1804→1805）。

【0119】

この結果、ユーザは、発言内容の表示中のアンカーを更にマウス装置等によってクリックすることにより、更に他のリソースにジャンプすることができる。

また、各行のデータが他の文書の行を引用したコメント行である場合には、表示装置306は、そのデータの色を変換するタグを追加した上で出力する（ステップ1806→1807）。

【0120】

この結果、ユーザは、コメント行を一目で判別することができる。

表示装置306は、該当する文書データの出力処理を終了すると、上記文書を含むスレッドのツリー構造を表示するHTML文書データを生成し出力して、発言内容表示の動作を終了する（ステップ1809→1810）。この処理は、前述した図14の動作フローチャートで示されるスレッドビューの表示動作と同様にして実現できる。

【0121】

以上の表示動作の結果、ユーザの端末のWebブラウザアプリケーション上には、例えば図28に示されるように、表示画面の上半分に発言内容が表示され、表示画面の下半分にはその発言内容の文書を含むスレッドのツリー構造が表示される。なお、この表示画面には、図28に示されるように、キーワードビューやスレッドビューを表示させるためのアンカーや、検索を実行するためのアンカーと等を同時に表示させることもできる。

【0122】

これらのビューの切替え機能により、例えば、キーワードビュー → スレッドビュー → 発言内容表示 → 発言者ビュー 発言内容表示 → スレッドビュー → … というように、会議室内の文書（発言）をユーザの嗜好に応じて横断的に参照してゆくことが可能となる。

【0123】

次に、表示装置306が実現する作者別／日付別色分け表示の動作につき説明する。図19は、その動作を示す動作フローチャートである。

まず、表示装置306は、メタインデックス304及びスレッドインデックス305に基づいて、図20(a)に示されるように作者項目毎に文書番号が分類された作者配列データと、図20(b)に示されるように日付項目毎に文書番号が分類された日付配列データとを予め作成する。これらの作成処理の詳細は省略するが、前述した図16の動作フローチャートと同様の処理によって実現できる。そして、作者配列データ中の各作者項目又は日付配列データ中の各日付項目に、それぞれ異なる色が割り当てられる。この色の割当ては、作者項目毎の作者の総数又は日付項目毎の日付の総数から決定される。

【0124】

次に、表示装置306は、ユーザの指定に基づく項目選択ボタン情報をWebサーバを経由して取得し、作者ボタンが押されたか日付選択ボタンが押されたかを判定する（ステップ1902、1904、図29参照）。

【0125】

表示装置306は、作者ボタンが押されたと判定した場合には、図20(a)に示される作者配列データを参照することにより、スレッドツリーの表示データを作成し出力する（ステップ1902→1903）。この処理は、前述した図14の動作フローチャートと同様の処理によって実現されるが、この場合に、ツリーの各ノードは、そのノードに対応する作者名に対応する作者配列データ中の作者項目に割り当てられている色で表示される。

【0126】

一方、表示装置306は、日付選択ボタンが押されたと判定した場合には、図20(b)に示される日付配列データを参照することにより、スレッドツリーの表示データを作成し出力する（ステップ1904→1905）。この処理も、前述した図14の動作フローチャートと同様の処理によって実現されるが、この場合に、ツリーの各ノードは、そのノードに対応する日付に対応する日付配列データ中の日付項目に割り当てられている色で表示される。

【0127】

以上の表示動作の結果、例えばユーザが作者ボタンを押した場合には、ユーザの端末のWebブラウザアプリケーション上には、例えば図29に示されるように、作者別に色分けされたスレッドのツリーが表示され、ユーザは同一の作者の文書を一目で確認することができる。

【0128】

次に、表示装置306が実現するスレッドビューを使った検索結果の強調表示の動作につき説明する。図21はその動作を示す動作フローチャートである。

まず、表示装置306は、検索後入力フォーム画面を表示するためのHTML文書データを生成し出力する（ステップ2101）。この結果、ユーザの端末のWebブラウザアプリケーション上には、例えば図30に示されるような検索入力フォーム画面が表示される。ユーザは、この検索入力フォームに検索語を入力して検索の実行を指定する。

【0129】

上記検索入力フォームに入力された検索語は、Webサーバを経由して文字列検索装置405（図4）に引き渡される。文字列検索装置405は、ユーザによる検索語の指定に基づいて、索引ファイル404を用いながら文書群データベース301内の指定されたスレッドを構成する各文書に対して全文検索を実行し、その検索語を含む文書番号を出力する（ステップ2102、2103）。

【0130】

表示装置306は、上記検索語を含む文書番号を受け取ると、その文書番号を含むスレッドのツリー構造を表示するHTML文書データを、前述した図14の動作フローチャートと同様の処理によって表示する。この場合に、表示装置306は、上記文書番号を含むノードの色を強調色に指定する（ステップ2104、2105）。

【0131】

この結果、ユーザの端末のWebブラウザアプリケーション上において、例えば図31に示されるような検索結果に基づくスレッドビューの強調表示が実現される。これにより、ユーザは、スレッドの構造を把握しつつ、検索を実行することができる。

【0132】

次に表示装置306が実現するキーワードビューを使った検索結果の強調表示の動作につき説明する。図22はその動作を示す動作フローチャートである。

まず、表示装置306は、図21のステップ2101の場合と同様に、検索後入力フォーム画面を表示するためのHTML文書データを生成し出力する（ステップ2201）。ユーザは、この検索入力フォームに検索語を入力して検索の実行を指定する。

【0133】

上記検索入力フォームに入力された検索語は、Webサーバを経由して文字列検索装置405（図4）に引き渡される。文字列検索装置405は、ユーザによる検索語の指定に基づいて、索引ファイル404を用いながら文書群データベース301内の指定された会議室を構成する各文書に対して全文検索を実行し、その検索語を含む文書番号を出力する（ステップ2202、2203）。

【0134】

表示装置306は、上記検索語を含む文書番号を受け取ると、まず、図6のデータ構造を有するスレッドインデックス305を参照して、上記文書番号を含むエントリに対応するルート文書番号を抽出する（ステップ2204）。

【0135】

続いて、表示装置306は、指定された会議室に関するキーワードビューを表示するHTML文書データを、前述した図13の動作フローチャートと同様の処理によって表示する。この場合に、表示装置306は、ステップ2204で抽出されたルート文書番号に対応するスレッドのタイトル又はその表示エリア全体の色を強調色に指定し、更に、表示されるキーワード中に検索語が含まれている場合には、そのキーワードも強調色に指定する（ステップ2205、2206、2207）。

【0136】

この結果、ユーザの端末のWebブラウザアプリケーション上において、例えば図32に示されるような検索結果に基づくキーワードビューの強調表示が実現される。これにより、ユーザは、検索語を含むスレッドを一目で把握することが

できる。

【0137】

なお、表示装置306は、検索結果の文書番号とそれに対応するタイトルを、例えば図33に示されるように羅列して表示するように構成することも可能である。

【0138】

最後に、表示装置306が実現するサブピック毎のキーワードビューの表示動作について説明する。

前述したキーワードビューは、スレッド毎にキーワードを表示するものであった。これに対して、サブピック毎のキーワードビューでは、1つのスレッド内のサブピック毎に、キーワードを抽出して表示することができる。

【0139】

この動作において、表示装置306は、図24のデータ構造を有するサブピックインデックスを使用する。サブピックインデックスは、図12に示されるカラーテーブルのデータ構造に対して、キーワードフィールドが追加されたデータ構造を有する。

【0140】

サブピックインデックスは、実質的には前述したカラーテーブルを置き換えるものであるため、サブピックインデックスにおけるキーワードフィールド以外のフィールドの内容は、構造解析部402による前述した図8のステップ810の処理によって予め登録されている。この場合、図8のステップ810の処理の説明において前述したように、カラーテーブルであるサブピックインデックスには、ルート文書番号毎（スレッド毎）に、それに含まれるルート文書のタイトルを示すメインピックと、それ以外の文書のタイトルを示すサブピックのそれぞれに対応するエントリが得られる。表示装置306は、この登録内容を利用する。

【0141】

図23は、表示装置306が実現するサブピックからのキーワード抽出の制御を示す動作フローチャートである。

まず、表示装置306は、各スレッドについて、サブトピックインデックス内のそのスレッドに含まれる各エントリに登録されている文書番号に基づいて、メイントピック及びサブトピック単位で、それぞれに属する文書群を各結合文書ファイルにまとめ（ステップ2301）、その結果得られる各結合文書ファイルを内容解析部403（図4）に入力する（ステップ2302）。

【0142】

内容解析部403は、各結合文書ファイル別にキーワードを抽出し、その結果を表示装置306に返す。表示装置306は、内容解析部403から返された各結合文書ファイル別のキーワードを、サブトピックインデックス内の上記各結合文書ファイルに対応するエントリのキーワードフィールドに登録する（ステップ2303）。

【0143】

以上のようにして、各スレッドについて、メイントピック及びサブトピック単位で、それぞれに属する文書群からキーワードが抽出される。

その後は、表示装置306は、サブトピックインデックスの内容に基づいて、ユーザにより指定されたスレッドに関して、そのスレッドのメイントピック及びサブトピック単位で、それぞれのタイトルとそれぞれに属するキーワードを表示するためのHTML文書データを生成し出力する。

【0144】

この結果、ユーザの端末のWebブラウザアプリケーション上には、例えば、図34に示されるような形式で、サブトピック毎のキーワードビューが表示される。これにより、ユーザは、キーワードによるより精密なトピックの絞り込みを行うことができる。

本実施の形態を実現するプログラムが記録された記録媒体についての補足

本発明は、計算機により使用されたときに、上述の本発明の実施の形態の各構成によって実現される機能と同様の機能を計算機に行わせるための計算機読出し可能記憶媒体として構成することもできる。

【0145】

この場合に、図35に示されるように、例えばフロッピーディスク、CD-R

OMディスク、光ディスク、リムーバブルハードディスク等の可搬型記憶媒体3502や、ネットワーク回線3503経由で、本発明の実施の形態の各種機能を実現するプログラムが、コンピュータ3501の本体3504内のメモリ（RAM又はハードディスク等）3505にロードされて、実行される。

【0146】

【発明の効果】

本発明の第1の態様の構成によれば、キーワードを頼りにして、雑多な文書集合の中から必要な文書が含まれているであろう文書群を正確かつ容易に見つけ出すことが可能となる。

【0147】

本発明の第2の態様の構成によれば、検索指定に該当する文書群を一目で把握することが可能となる。

本発明の第3の態様の構成によれば、文書群全体の構造を容易に把握することが可能となる。

【0148】

本発明の第4の態様の構成によれば、文書群全体の話題の推移を容易に把握することが可能となる。

本発明の第5の態様の構成によれば、スレッドの構造を把握しつつ、検索を実行することが可能となる。

【0149】

本発明の第6の態様の構成によれば、キーワードによるより精密なトピックの絞込みを行うことが可能となる。

このように、本発明によれば、文書群に対して種々の観点からアクセスすることが可能となる。

【図面の簡単な説明】

【図1】

参照関係を有する文書群の例を示す図である。

【図2】

文書のデータ構造の例を示す図である。

【図 3】

本発明の実施の形態のシステム構成図（その 1）である。

【図 4】

本発明の実施の形態のシステム構成図（その 2）である。

【図 5】

メタインデックスのデータ構造を示す図である。

【図 6】

スレッドインデックスのデータ構造を示す図である。

【図 7】

索引ファイルのデータ構造を示す図である。

【図 8】

書式解析部と構造解析部の動作フローチャートである。

【図 9】

文書番号のスレッドインデックスへの登録の動作フローチャートである。

【図 10】

文書番号のスレッドインデックスへの登録の動作説明図である。

【図 11】

色番号登録の動作フローチャートである。

【図 12】

カラーテーブルの例を示す図である。

【図 13】

キーワードビューの動作フローチャートである。

【図 14】

スレッドビューの動作フローチャートである。

【図 15】

スレッドビューの制御用配列の例を示す図である。

【図 16】

発言者ビューの動作フローチャートである。

【図 17】

発言者ビューの制御用配列の例を示す図である。

【図18】

発言内容表示の動作フローチャートである。

【図19】

作者別／日付別色分け表示の動作フローチャートである。

【図20】

作者別／日付別色分け表示用配列の例を示す図である。

【図21】

スレッドビューを使った検索結果の強調表示の動作フローチャートである。

【図22】

キーワードビューを使った検索結果の強調表示の動作フローチャートである。

【図23】

サブピックからのキーワード抽出の制御動作フローチャートである。

【図24】

サブピックインデックスの例を示す図である。

【図25】

キーワードビューの表示例を示す図である。

【図26】

スレッドビューの表示例を示す図である。

【図27】

発言者ビューの表示例を示す図である。

【図28】

発言内容の表示例を示す図である。

【図29】

スレッドビューを用いた文書属性「作者」の強調（色別）表示の例を示す図である。

【図30】

会議室内の全発言の検索（入力）表示例を示す図である。

【図31】

スレッドビューを用いた文字列「プロトコル」を含むノードの強調表示の例を示す図である。

【図32】

キーワードビューを用いた文字列「プロトコル」を含むスレッドの強調表示の例を示す図である。

【図33】

会議室内の全発言の検索（結果出力）表示例を示す図である。

【図34】

サブトピックから抽出したキーワードの表示例を示す図である。

【図35】

本実施の形態を実現するプログラムが記録された記録媒体の説明図である。

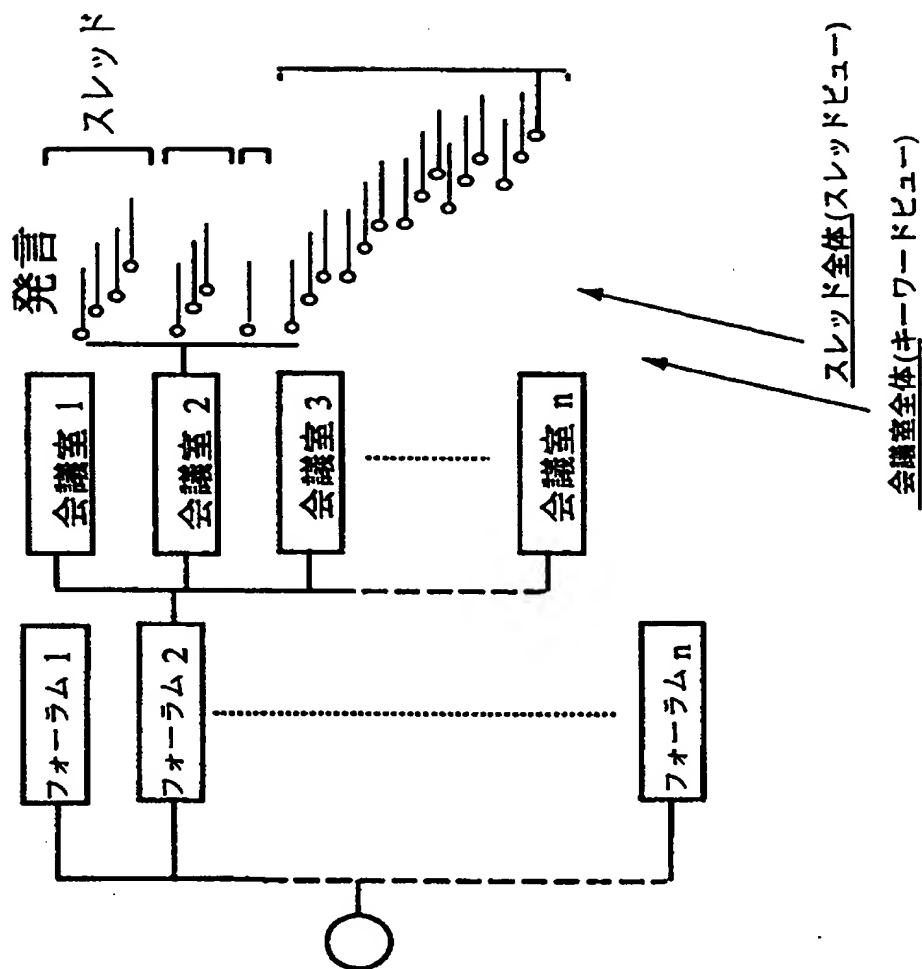
【符号の説明】

301	文書群データベース
302	文書群解析装置
303	集計装置
304	メタインデックス
305	スレッドインデックス
306	表示装置
401	書式解析部
402	構造解析部
403	内容解析部
404	索引ファイル
405	文字列検索装置

【書類名】 図面

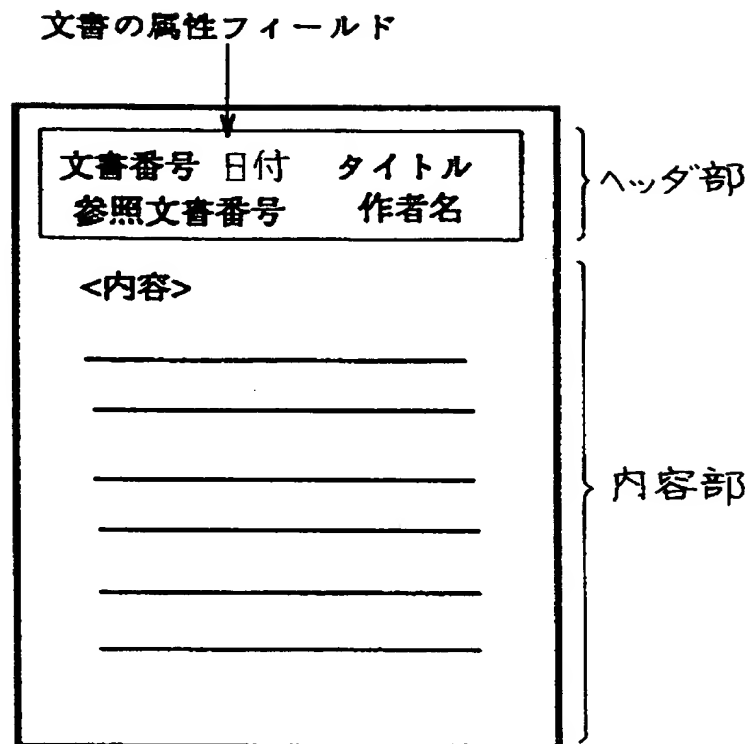
【図1】

参照関係を有する文書群の例を示す図



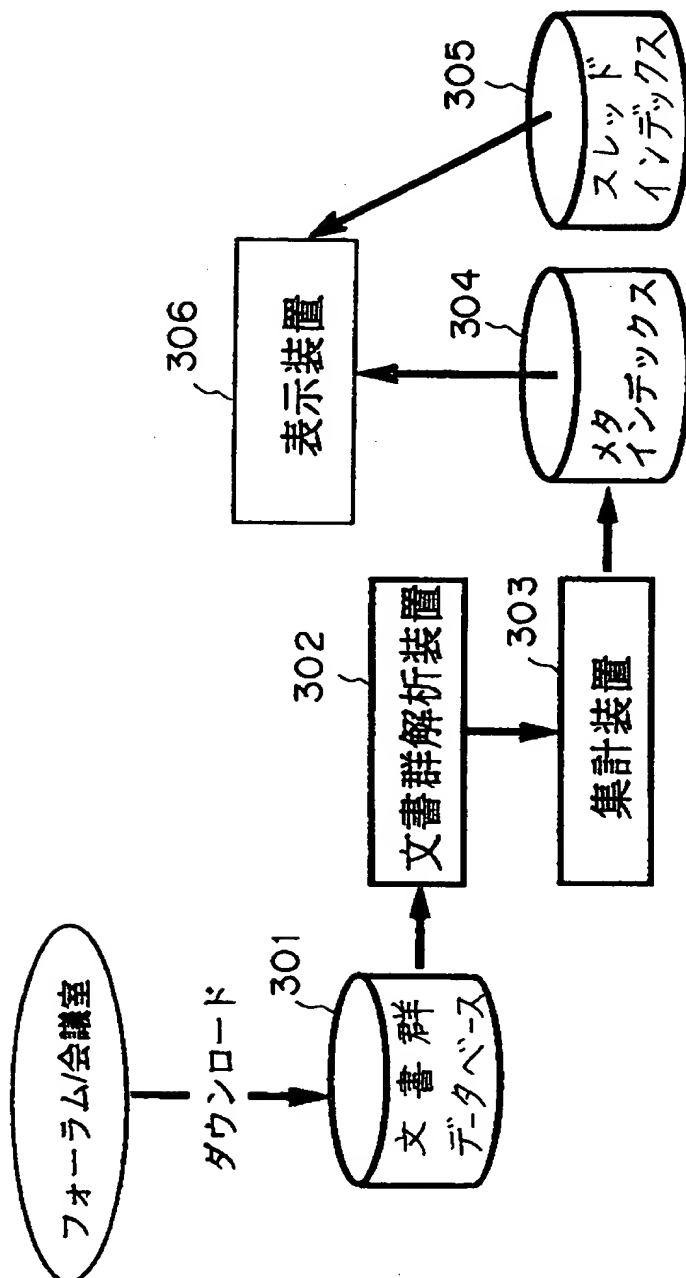
【図2】

文書のデータ構造の例を示す図



【図3】

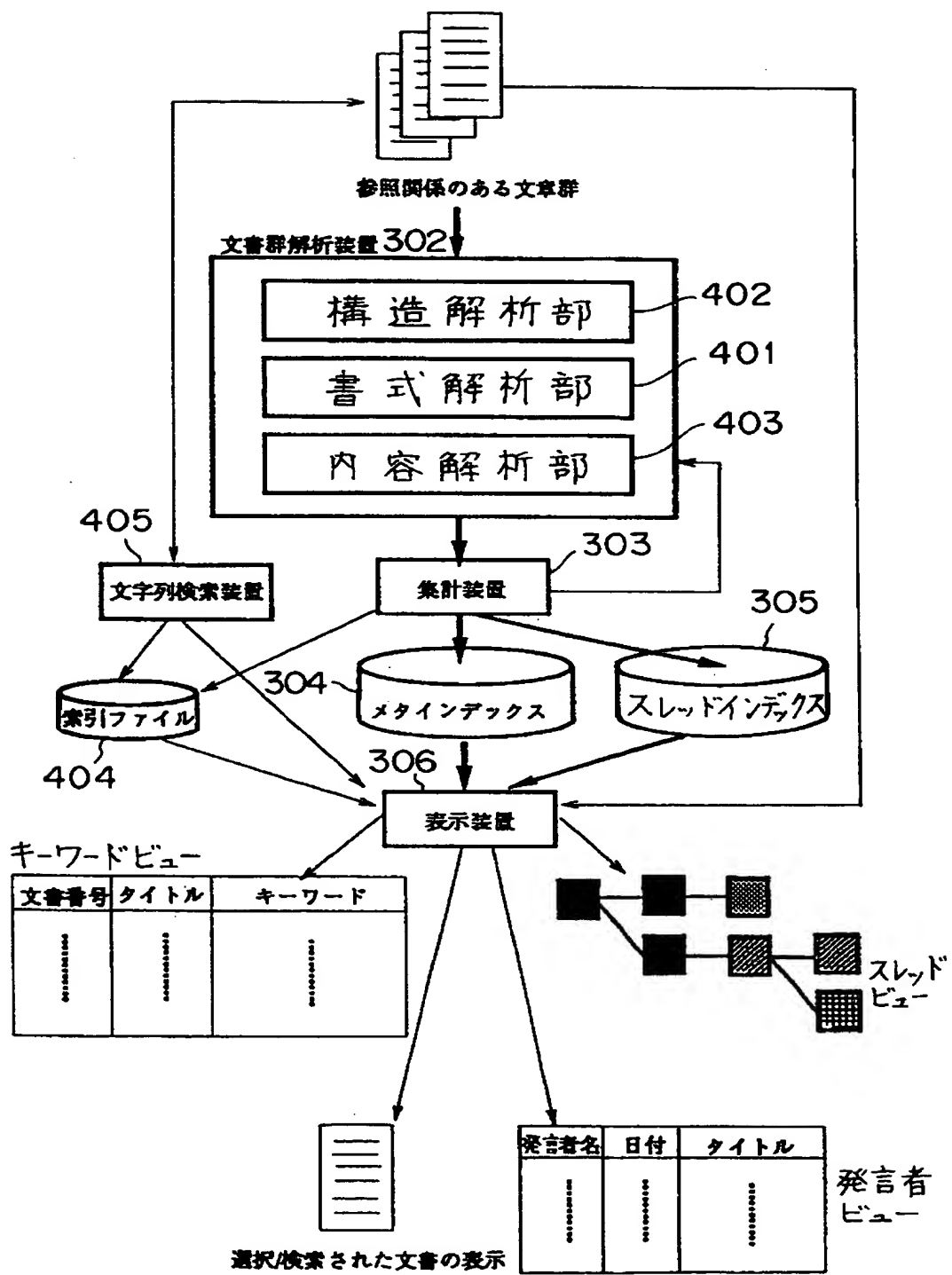
本発明の実施の形態のシステム構成図
(その1)



特平 9-242247

【図4】

本発明の実施の形態のシステム構成図
(その2)



【図5】

メタインデックスのデータ構造を示す図

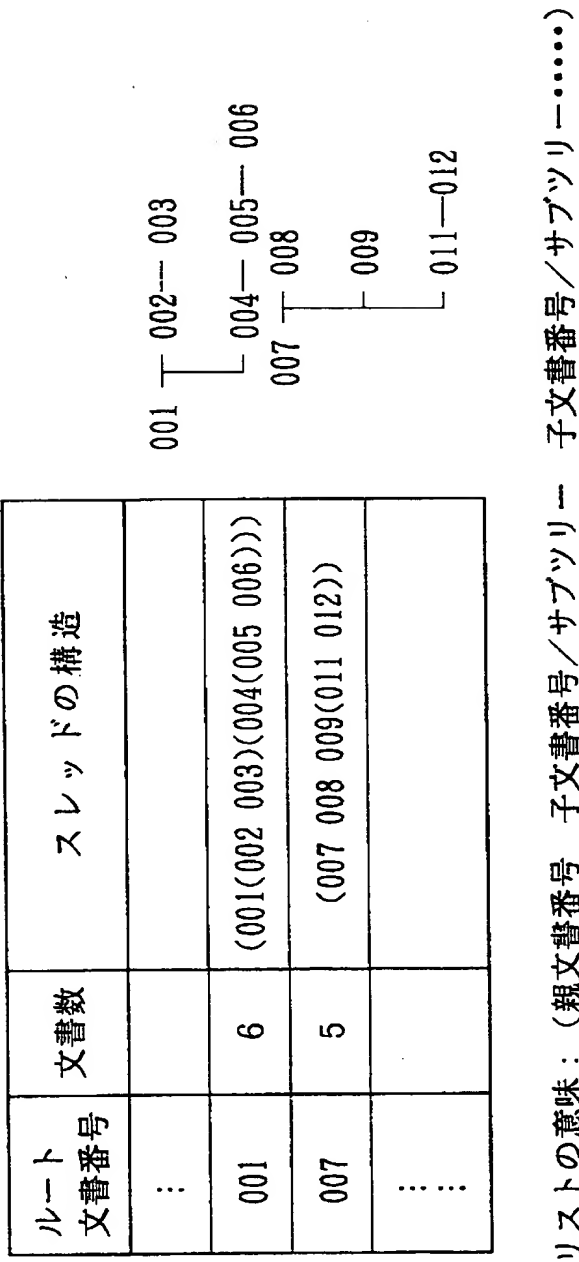
文書番号	参照文書番号	タイトル	作者名	キーワード	行数	色番号	日付
1	1	1	1	1	1	1	1
004	001	教えて	内野	—	12	4	96/10/05
005	004	Re:教えて	津田	—	45	4	96/10/08
006	005	Javaのアプリ	内野	—	80	6	96/10/15
*007	—	HD不良?	松井	SunOS,HD,64MB,フォーマット	12	1	96/10/18
-----	-----	-----	-----	-----	-----	-----	-----

*はルートの文書番号

特平 9-242247

【図6】

スレッドインデックスのデータ構造を示す図



【図7】

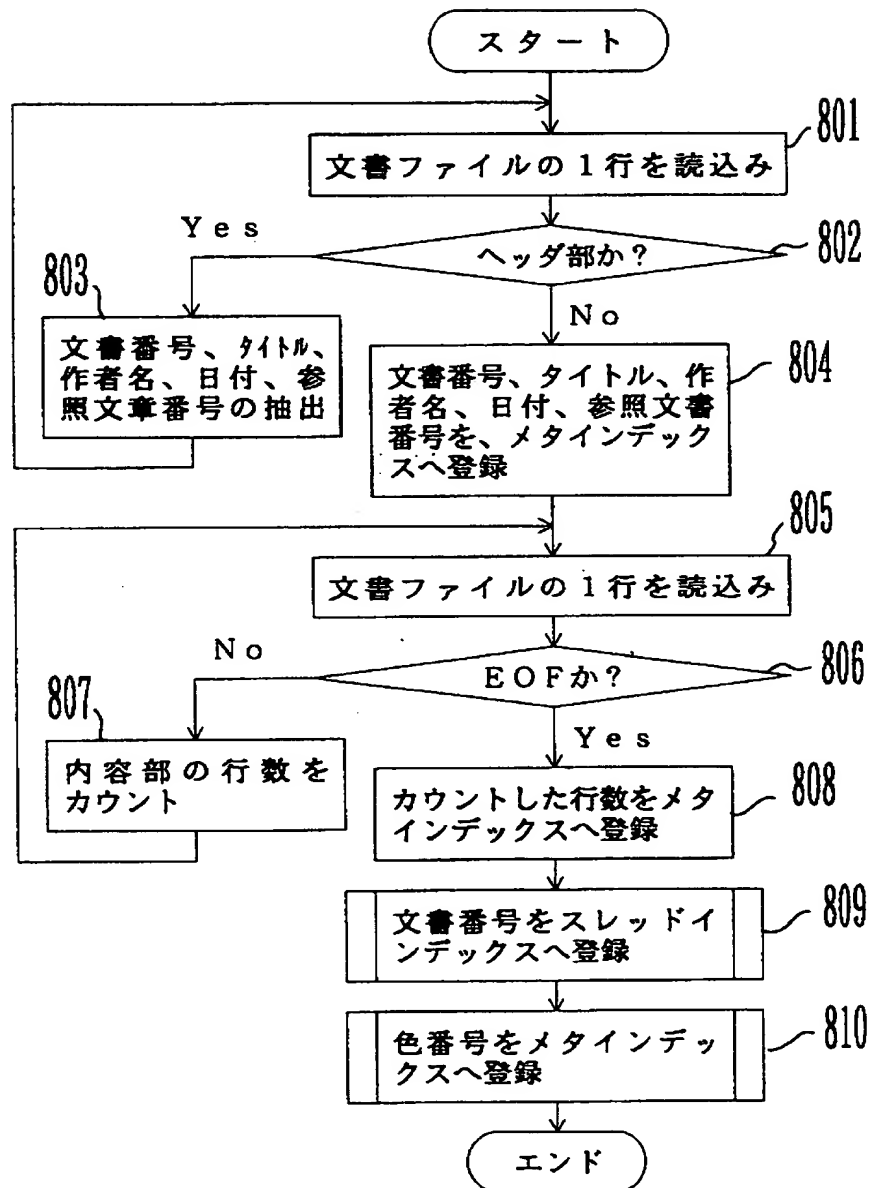
索引ファイルのデータ構造を示す図

索引語	索引語を含むキーワード	スレッド番号*
インターネット	・インターネットスタート ・インターネットへの接続	00496 00218
コントロール	・コントロールパネル	00518
ドメイン	・ドメインネームサーバ ・国のドメイン	00128 00476
⋮	⋮	⋮

*スレッド番号はスレッドのルートの文書番号

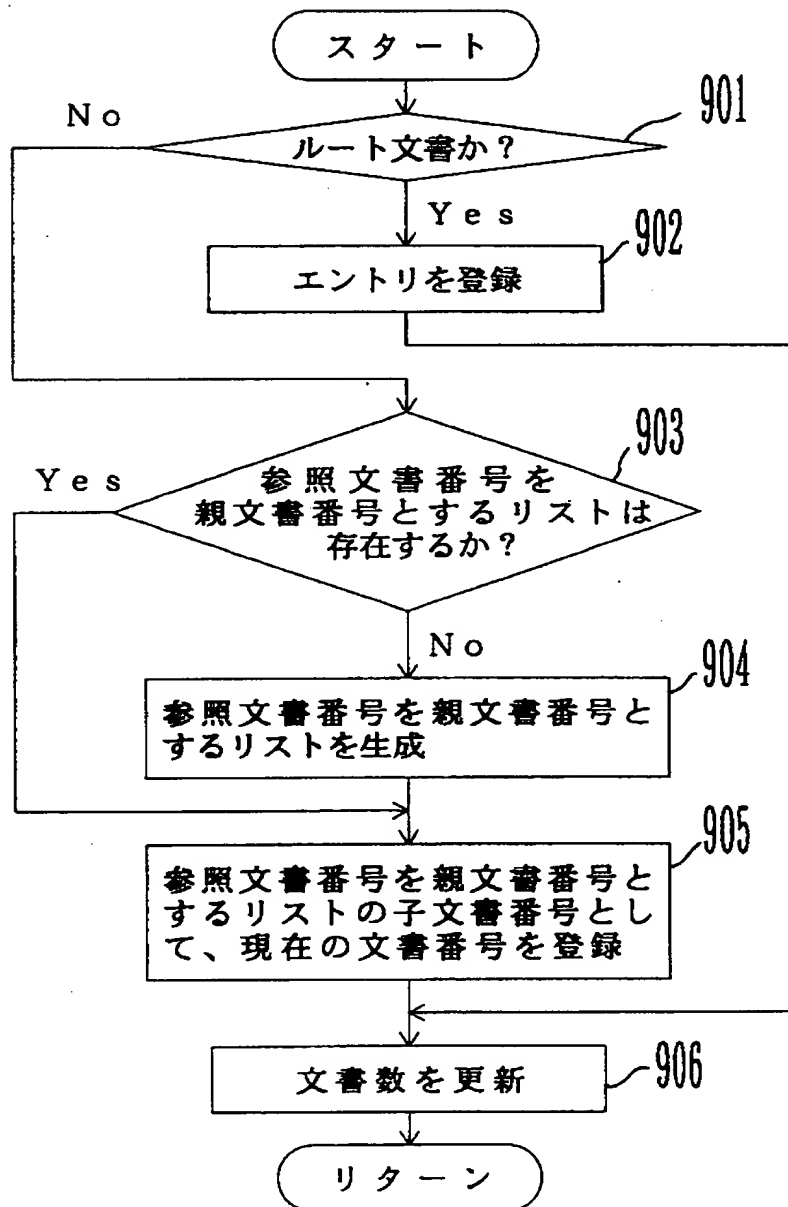
【図 8】

書式解析部と構造解析部の動作フローチャート



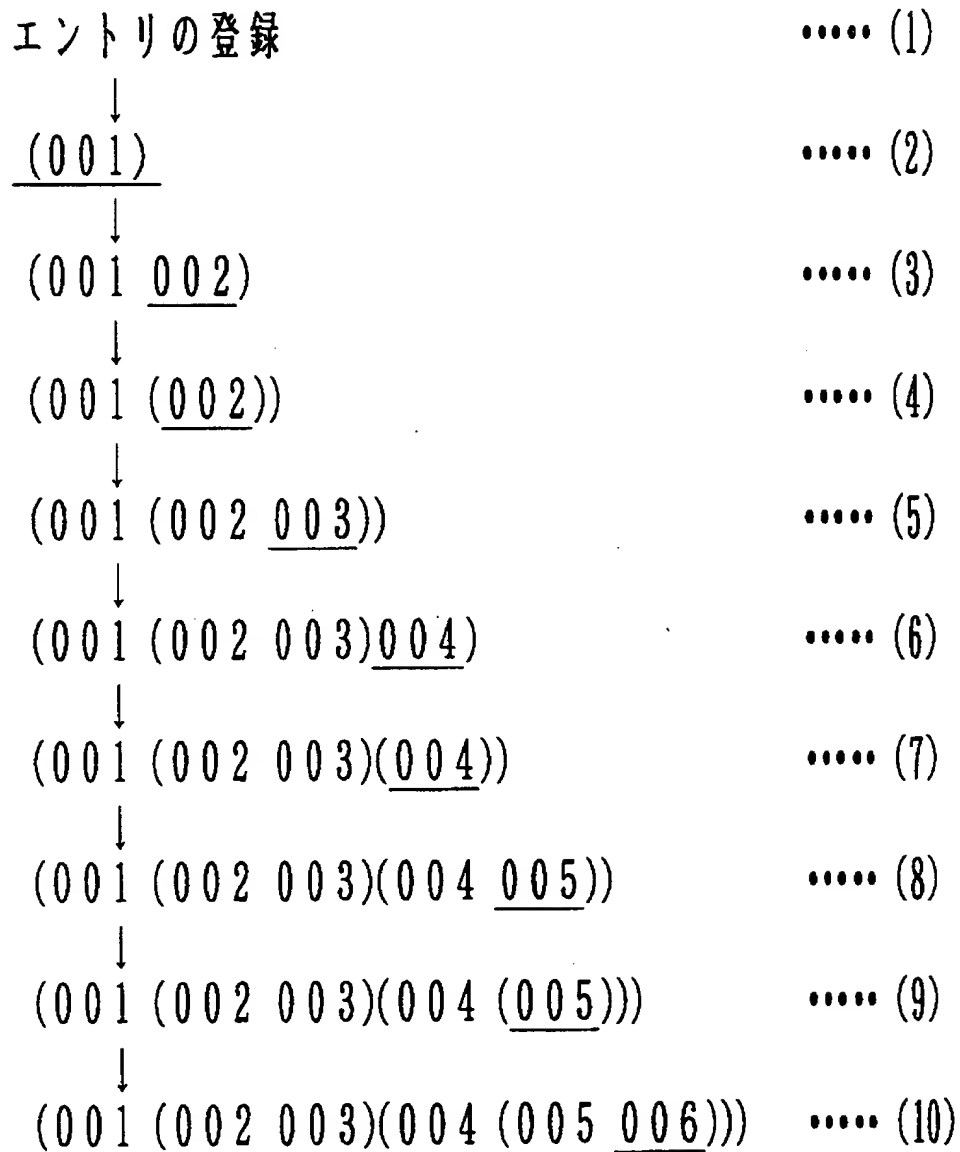
【図9】

文書番号のヘッドインデックスへの登録の動作フローチャート



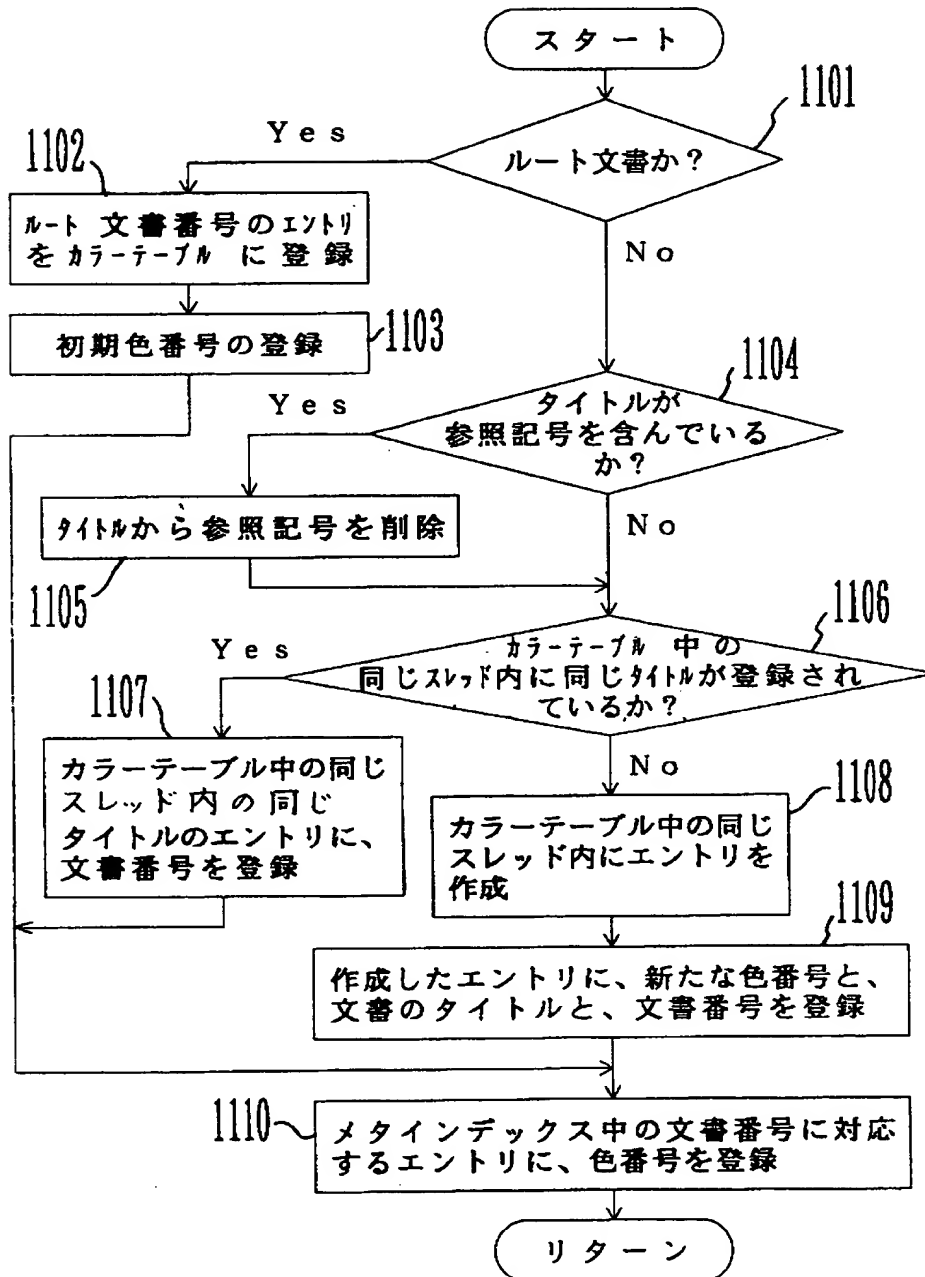
【図10】

文書番号のスレッドインデックスへの登録の動作説明図



【図11】

色番号登録の動作フローチャート



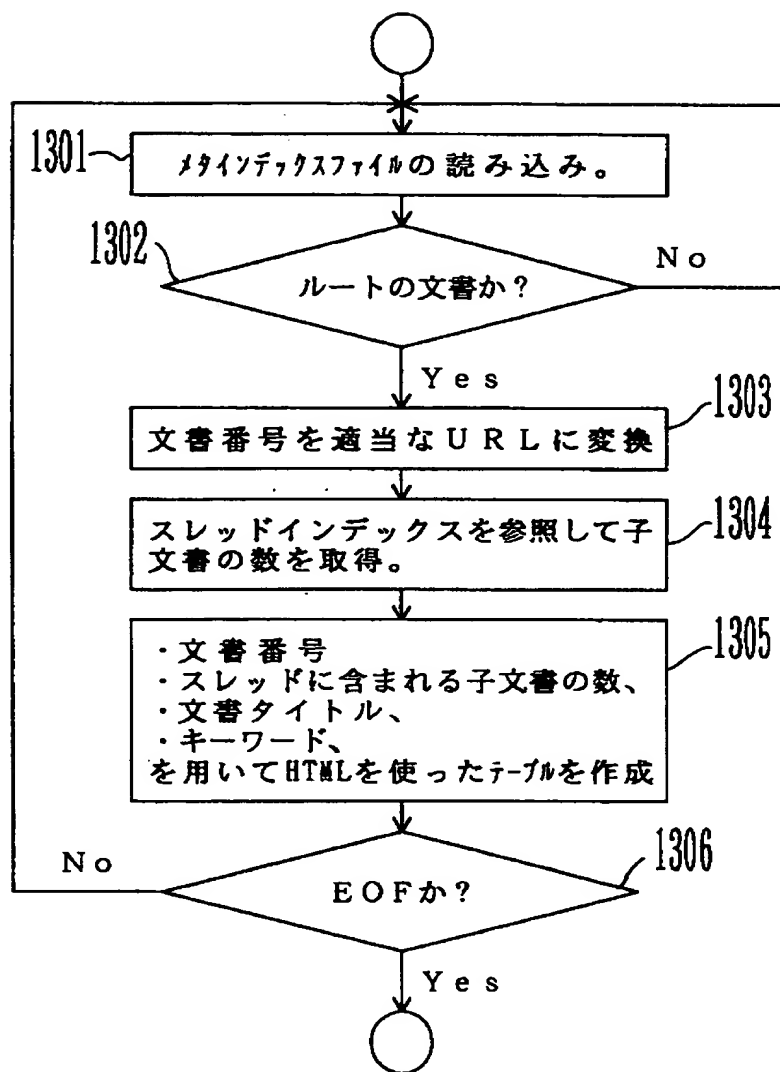
【図12】

カラーテーブルの例を示す図

ル-ト 文書番号	色番号	タイトル	文書番号
⋮	⋮	⋮	⋮
001	#1	(メインピク)	001, 002, 004
	#2	(サブピク 1)	003
	#3	(サブピク 2)	005, 006
⋮	⋮	⋮	⋮

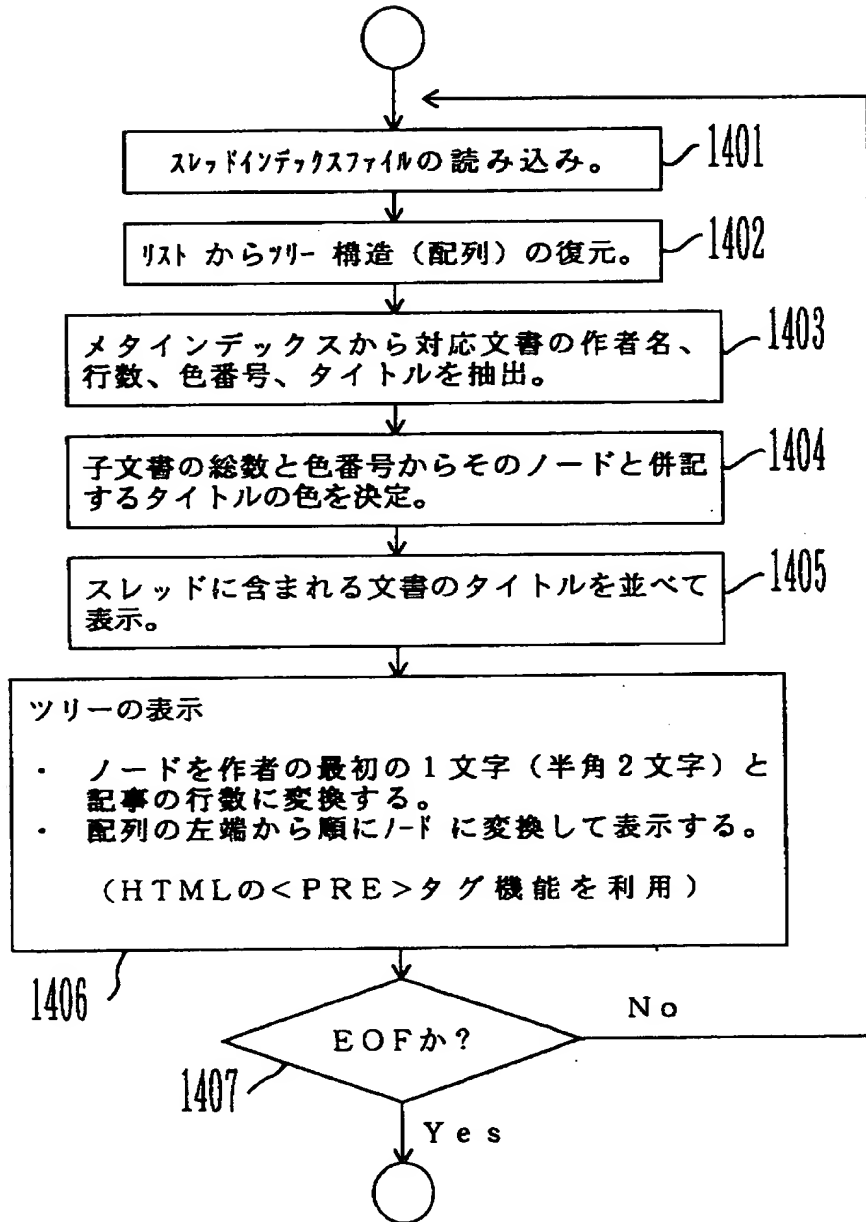
【図13】

キーワードビューの動作フローチャート



【図14】

スレッドビューの動作フローチャート



【図15】

スレッドビューの制御用配列の例を示す図

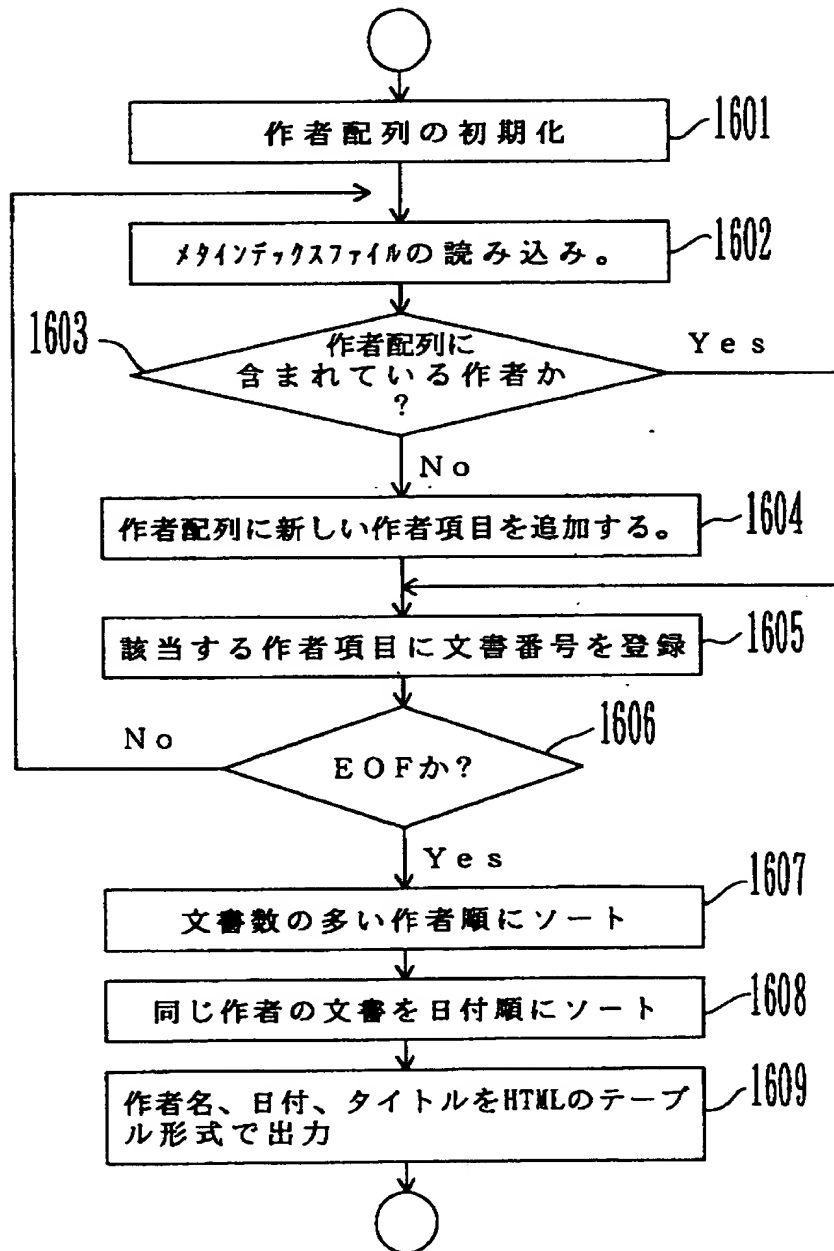
ツリー構造のリスト：(001(002 003)(004(005 006)))

	1	2	3	4
1	001	002	003	空白
2	空白	004	005	006

ツリー構造を表す配列

【図16】

発信者ビューの動作フローチャート



【図 17】

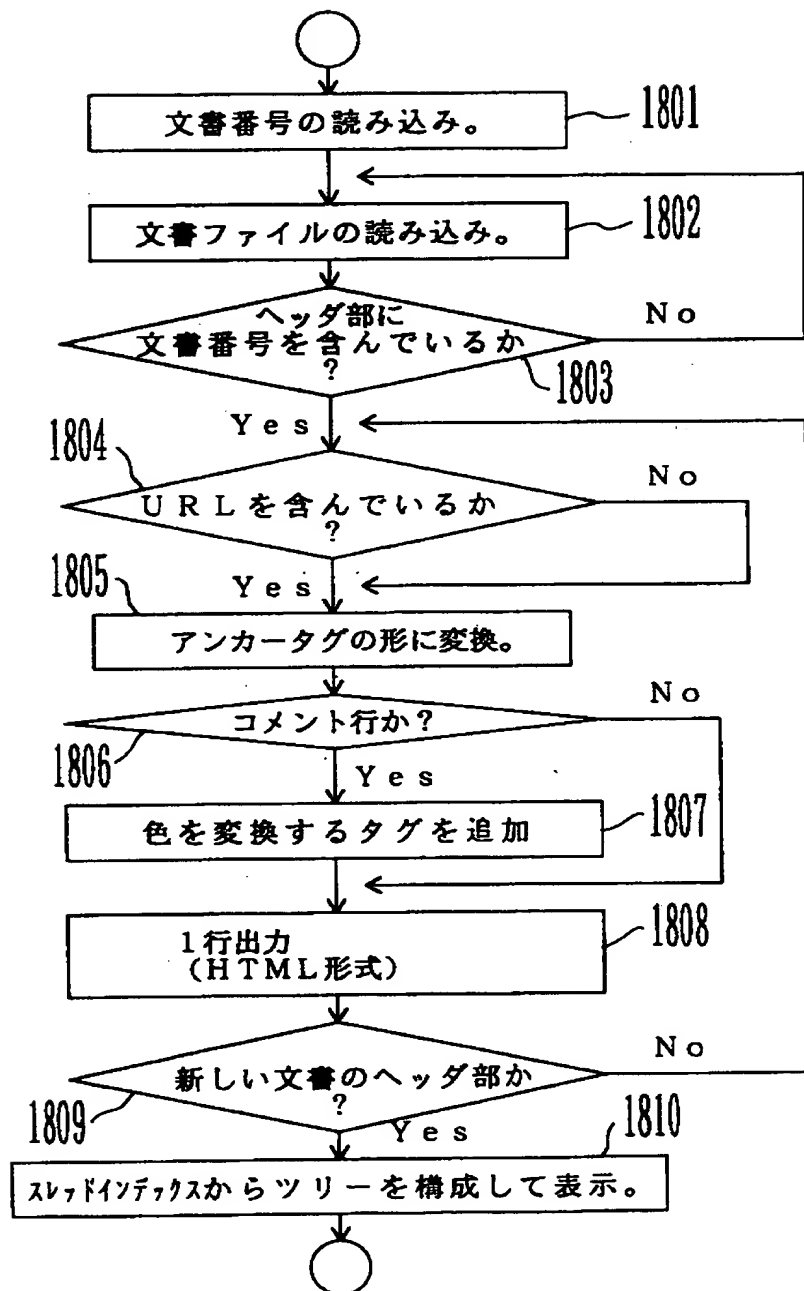
発言者ビューの制御用配列の例を示す図

作者配列の例

作者 A	001, 005, 007, 008
作者 B	011, 012, 013
作者 C	003, 009, 010
作者 D	004, 006
作者 E	002

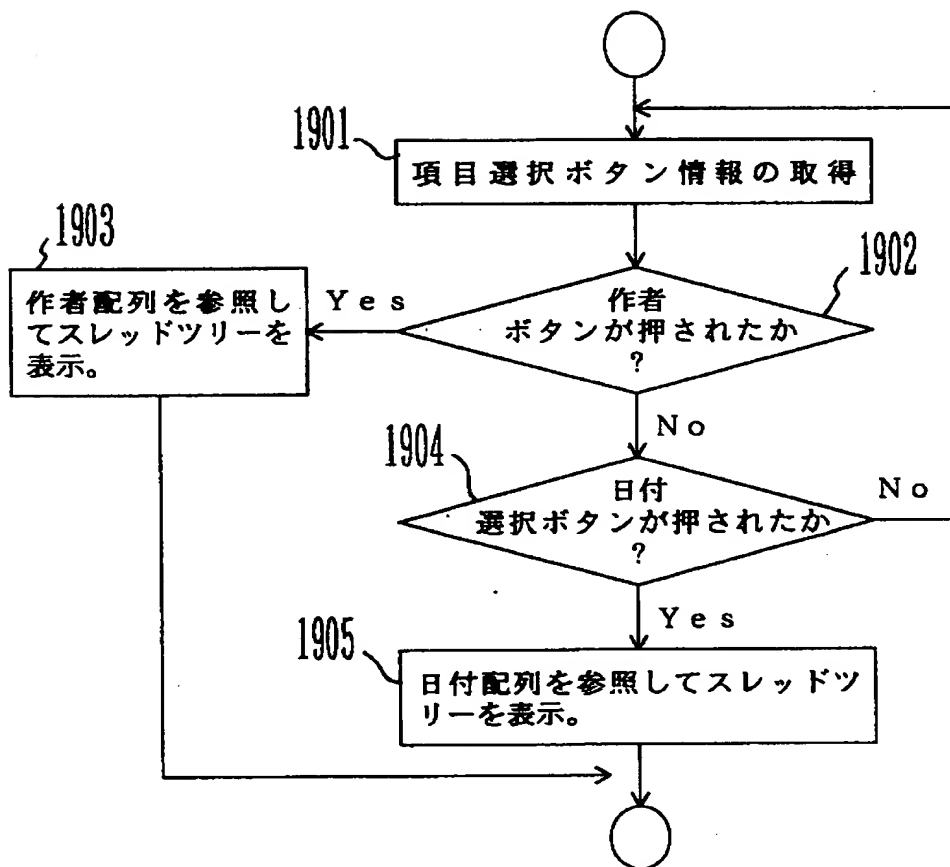
【図18】

発言内容表示の動作フローチャート



【図19】

作者別／日付別色分け表示の動作フローチャート



【図20】

作者別 日付別
色分け表示用配列の例を示す図

(b) 日付配列の例

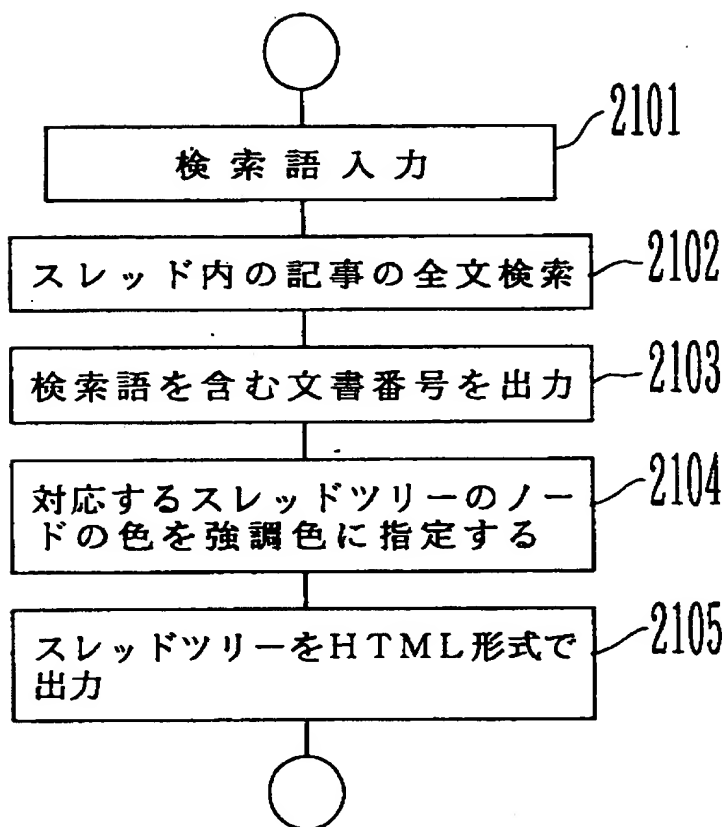
97/04/15	001, 002, 003, 004	↑	黄
97/04/21	005, 006	↑	黒
97/04/28	007, 008, 009	↑	赤
97/05/25	010	↑	青
97/05/26	011, 012, 013, 014, 015	↑	緑

(a) 作者配列の例

作者A	001, 005, 007, 008	↑	黄
作者B	011, 012, 013	↑	黒
作者C	003, 009, 010	↑	赤
作者D	004, 006	↑	青
作者E	002	↑	緑

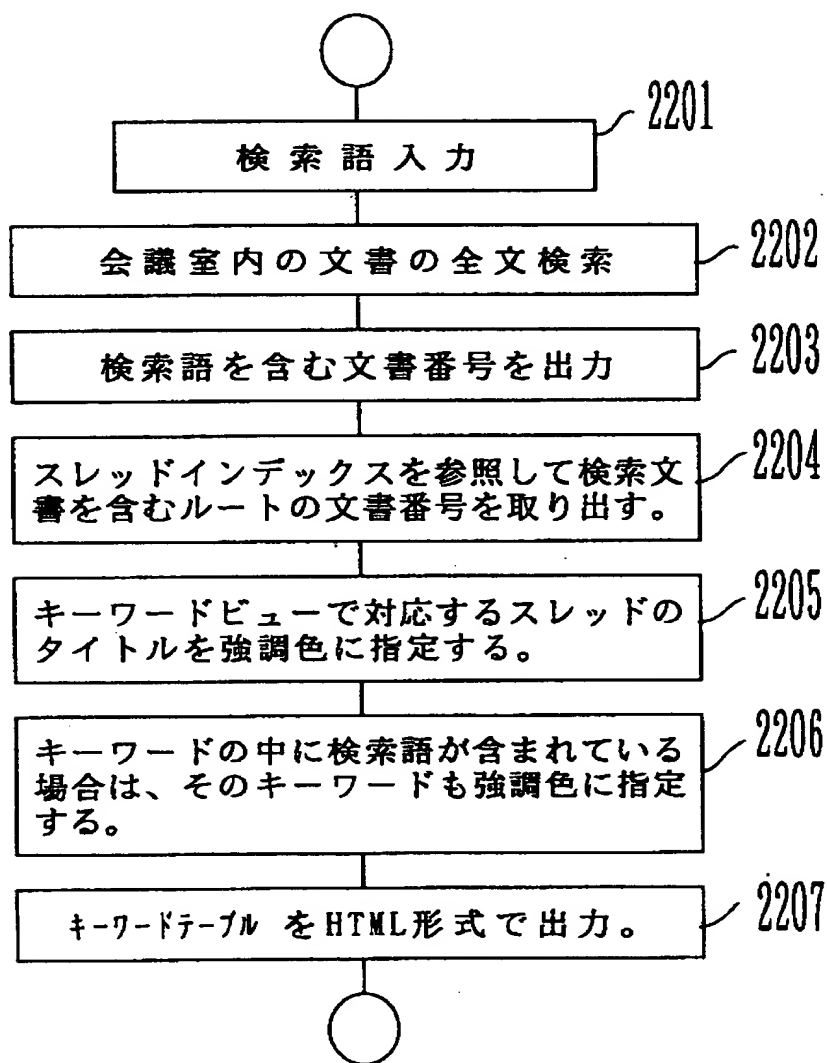
【図21】

スレッドビューを使った検索結果の 強調表示の動作フローチャート



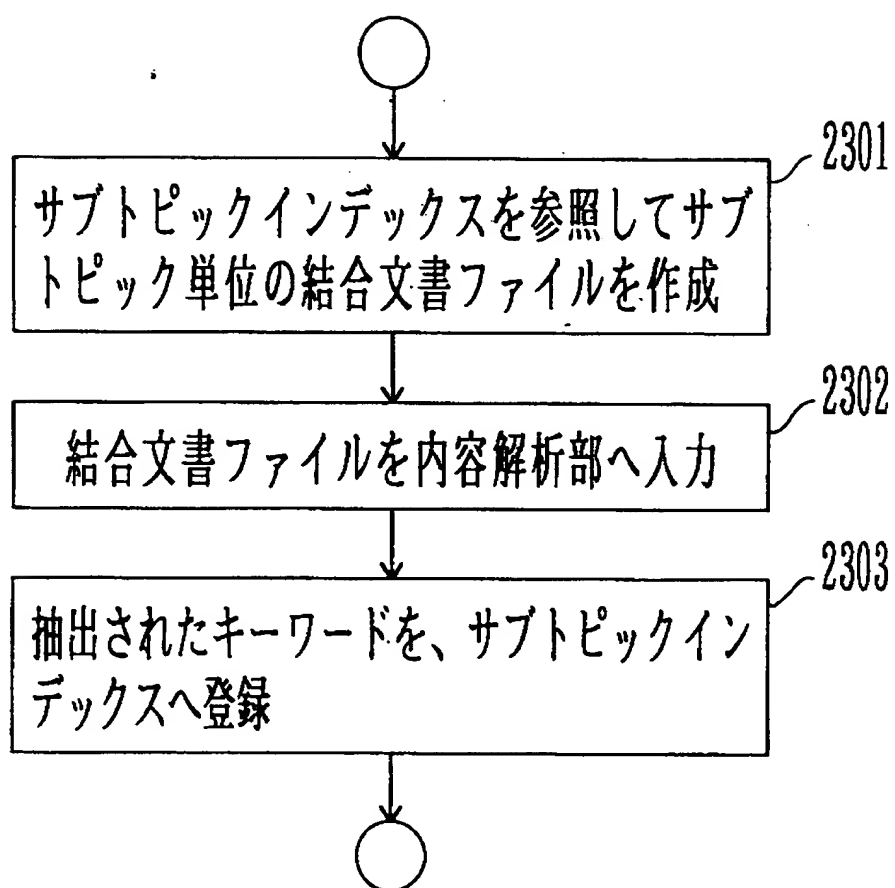
【図22】

キーワードビューを使った検索結果の
強調表示の動作フローチャート



【図23】

サブトピックからのキーワード抽出の
制御動作フローチャート



【図24】

サブピックインデックスの例を示す図

主文書番号	色番号	タイトル	記事番号	キーワード
∴	∴	∴	∴	∴
001	#1	(メインピック)	001, 002, 004	∴∴∴∴
	#2	(サブピック1)	003	∴∴∴∴
	#3	(サブピック2)	005, 006	∴∴∴∴
∴	∴	∴	∴	∴

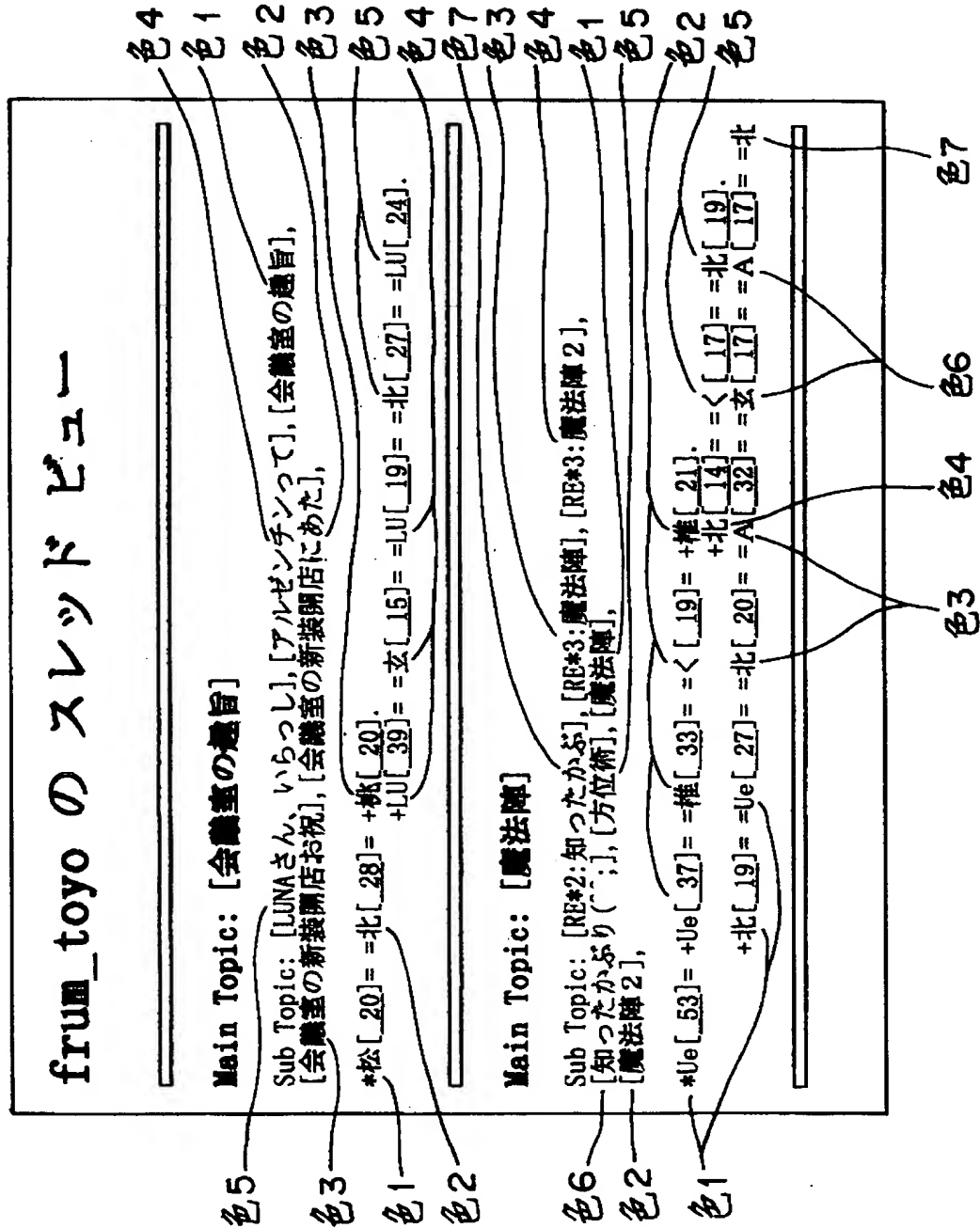
【図25】

キーワードビューの表示例を示す図

ID	タイトル	キーワード
00001	1 はじめに	質問の内容 回答 接続形態 用環境(ハードウェア) 的確な内容の アドバイス はば/稲葉 Manager Internet 関連の
00002	3 最低限のハードウェア	最近流行のネットサーフィン インターネット デファクト・スタンダード のネットスターブ ナビゲータ ペンデ・アムマシン こっち ならパソ通 法保の条文 色んなサービス PeopleのWorldTalk MIDI
00005	3 質問: ホームページにつながらない	プロバイダ パスワードの確認 使用形態 設定 ネットケーブルのデフォルト デイヤルアップ 接続 使用ソフト FMV*DPSP5 ユーザ名
00008	5 インターネット	プロバイダ ホームページ 料金 テレホーダイの時間開始 www アクセス ISDN モデム 28.8 回線速度 nifty のコマンド
00014	1 皆さん有り難う	リーヤさん推薦のプロバイダ 勝手気便アバウトネットワーカー 明日電話 金の社員を回! ま テレホー いろいろなチャレンジ せ名古屋 ROAD5 この間 子会社
00015	4 ニフティのダイヤルアップ接続	niftyserve つけ スクリプト SubManager or 会星人 NIFTYServe インターネット ウィザード kazusi
00016	1 プロバイダ選び	勝手気便アバウトネットワーカー オンラインサイ リーヤさん推薦 ホームページの設定 資料請求 ソファップ 固定料金 程度信頼 結構痛 ALLESNET
00030	9 プロバイダー 一月 自宅PC 間の速度	サーバーサーバー自宅PC niftyserve プロキシ I*E ping MNP5 下の方のレイヤのキャッシュ SubManager 会星人 モデムコマンド
00040	ニフ接続のメール	インターネットのメール niftyserve SubManager or はば/稲葉 会星人 ニフ接続の人 NIFTYServe サービスの中 kazusi

【図26】

スレッドビューの表示例を示す図



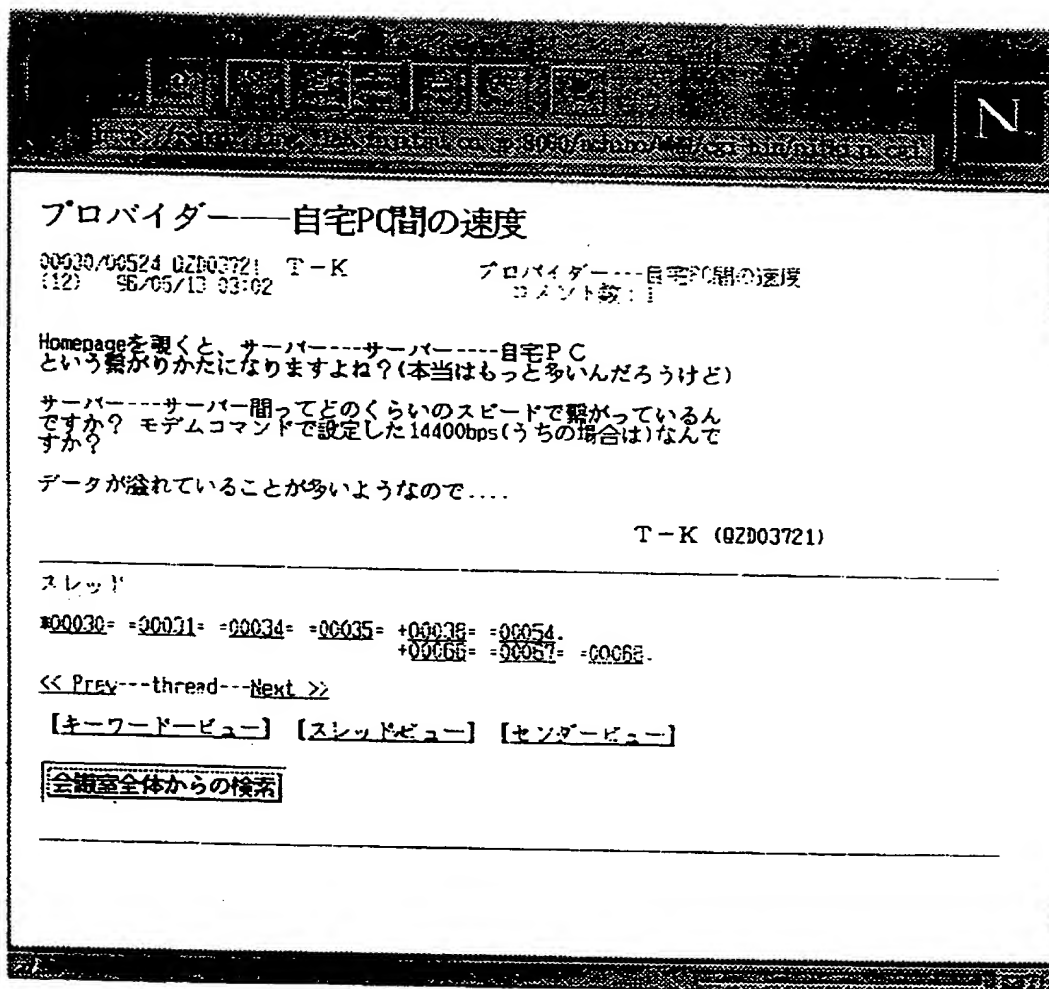
【図27】

発言者ビューの表示例を示す図

		8日: [RE:4:SONYNETをこぞにの方おられます?]
MURA	9月	11日: [RE:NTのAdmin権限] 30日: [ネスケ3.0i86でQuickTime]
	10月	1日: [RE:2:ネスケ3.0i86でQuickTime] [RE>NNJJa8のインストール出来ず] 2日: [RE>ネスケ3.0i86でQuickTime] [RE>ネスケ3.0i86でQuickTime]
	7月	25日: [98+DOSでnewsを読みたい] 27日: [RE:98+DOSでnewsを読みたい]
かんちゃん ★	9月	29日: [拡張子 rarについて] [RE:匿名メール] 30日: [RE:拡張子 rarについて]
	10月	1日: [RE:拡張子 rarについて]
	5月	8日: [RE:2:ニフティのダイヤルアップ接続] 19日: [RE:5:ニフティのダイヤルアップ接続]
よしちゅ	8月	5日: [RE:2:SONYNETをこぞにの方おられます?] 6日: [RE:3:SONYNETをこぞにの方おられます?]
	8月	3日: [ActiveXの初歩的質問です] 6日: [RE:ActiveXの初歩的質問です] 31日: [IE3正式版トラブル]
	9月	2日: [RE:IE3正式版トラブル] 3日: [RE:IE3正式版トラブル]
荒木 修	9月	9日: [メールに繋がらない(2)] 14日: [RE:メールに繋がらない(2)] 15日: [RE:メールに繋がらない(2)] 15日: [RE:一般的な接続方法]
	9月	9日: [メールに繋がらない(2)]
	9月	14日: [RE:メールに繋がらない(2)] 15日: [RE:メールに繋がらない(2)] 15日: [RE:一般的な接続方法]
	9月	15日: [RE:一般的な接続方法]
酒匂 歳弘	9月	9日: [メールに繋がらない(2)] 14日: [RE:メールに繋がらない(2)] 15日: [RE:メールに繋がらない(2)] 15日: [RE:一般的な接続方法]
	9月	9日: [メールに繋がらない(2)]
	9月	14日: [RE:メールに繋がらない(2)] 15日: [RE:メールに繋がらない(2)] 15日: [RE:一般的な接続方法]
Mue	8月	4日: [RE: (質問) キーボードの印刷はどうやって] 19日: [RE:Hyper RoadとRoad 4でのPPPPPP] 25日: [RE:INET SETUP WIZARDについて]
	8月	4日: [RE: (質問) キーボードの印刷はどうやって] 19日: [RE:Hyper RoadとRoad 4でのPPPPPP] 25日: [RE:INET SETUP WIZARDについて]
	8月	4日: [RE: (質問) キーボードの印刷はどうやって] 19日: [RE:Hyper RoadとRoad 4でのPPPPPP] 25日: [RE:INET SETUP WIZARDについて]
安藤貞之	7月	27日: [トラぶってしまいました SOS!] 29日: [RE:トラぶってしまいました SOS!] 30日: [RE:トラぶってしまいました SOS!]
	7月	27日: [トラぶってしまいました SOS!] 29日: [RE:トラぶってしまいました SOS!] 30日: [RE:トラぶってしまいました SOS!]
	7月	27日: [トラぶってしまいました SOS!] 29日: [RE:トラぶってしまいました SOS!] 30日: [RE:トラぶってしまいました SOS!]
N-DATE	8月	19日: [Real Audioのプラグインについて] 19日: [RE:Real Audioのプラグインについて] 19日: [RE:Real Audioのプラグインについて]
	8月	19日: [Real Audioのプラグインについて] 19日: [RE:Real Audioのプラグインについて] 19日: [RE:Real Audioのプラグインについて]
	8月	19日: [Real Audioのプラグインについて] 19日: [RE:Real Audioのプラグインについて] 19日: [RE:Real Audioのプラグインについて]

【図28】

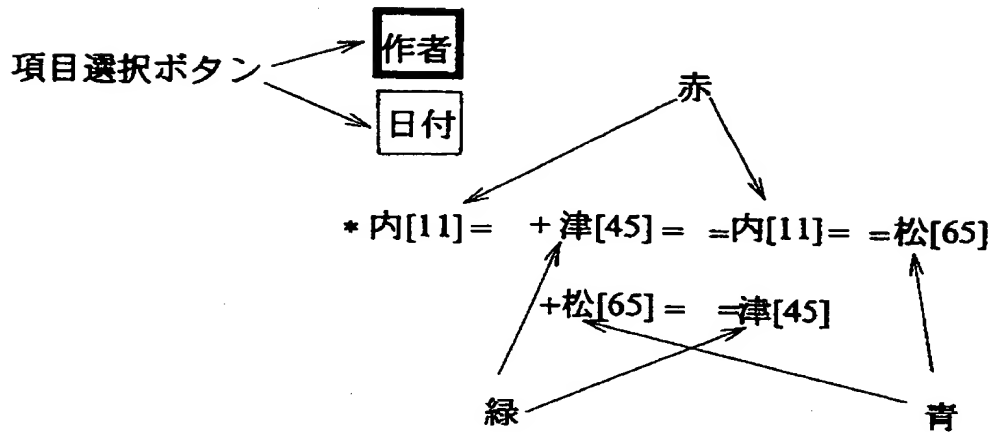
発言内容の表示例を示す図



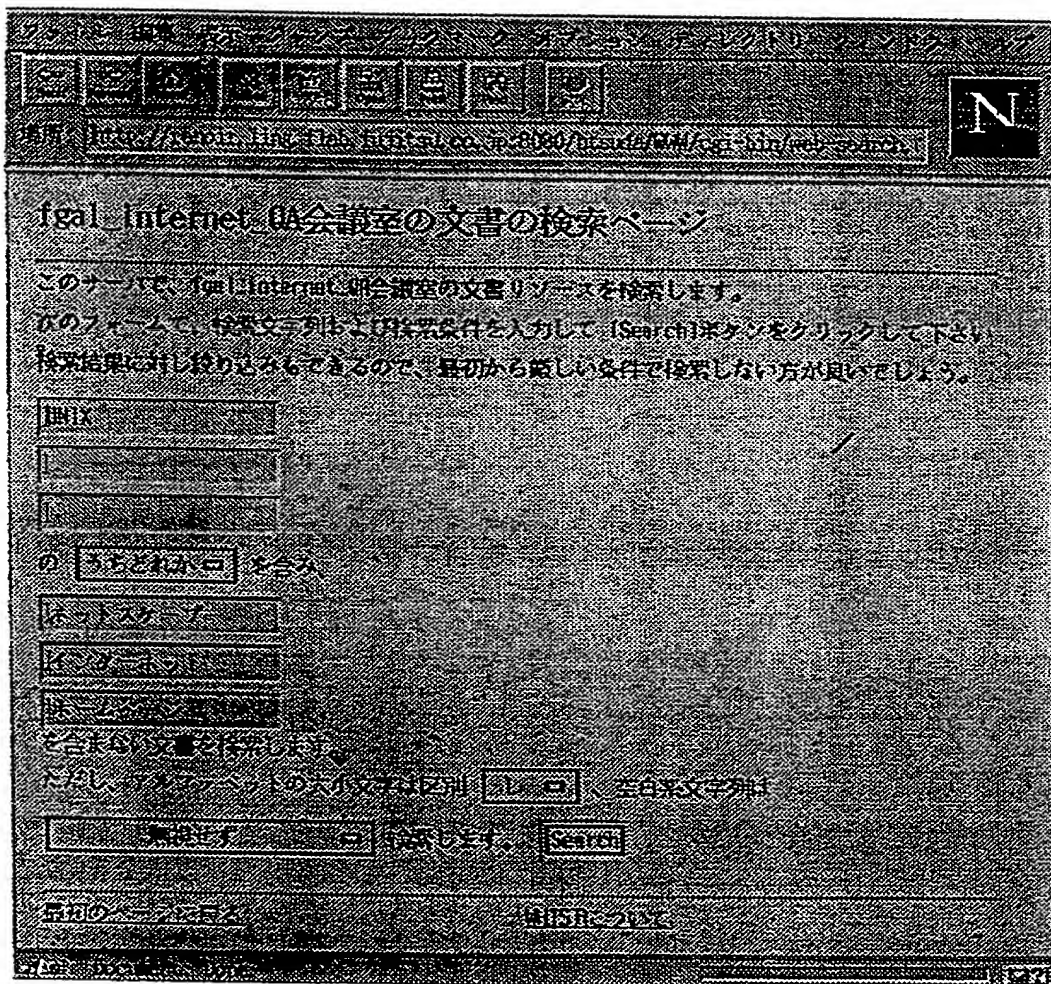
【図29】

スレッドビューを用いた文書属性

「作者」の強調(色別)表示の例を示す図



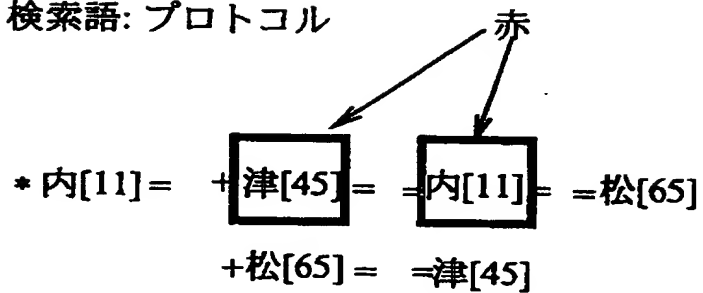
会議室内の全発言の検索(入力)表示例を
示す図



【図31】

スレッドビューを用いた文字列
「プロトコル」を含むノードの強調表示の
例を示す図

検索語: プロトコル



【図32】

キーワードビューを用いた文字列
「プロトコル」を含むスレッドの
強調表示の例を示す図

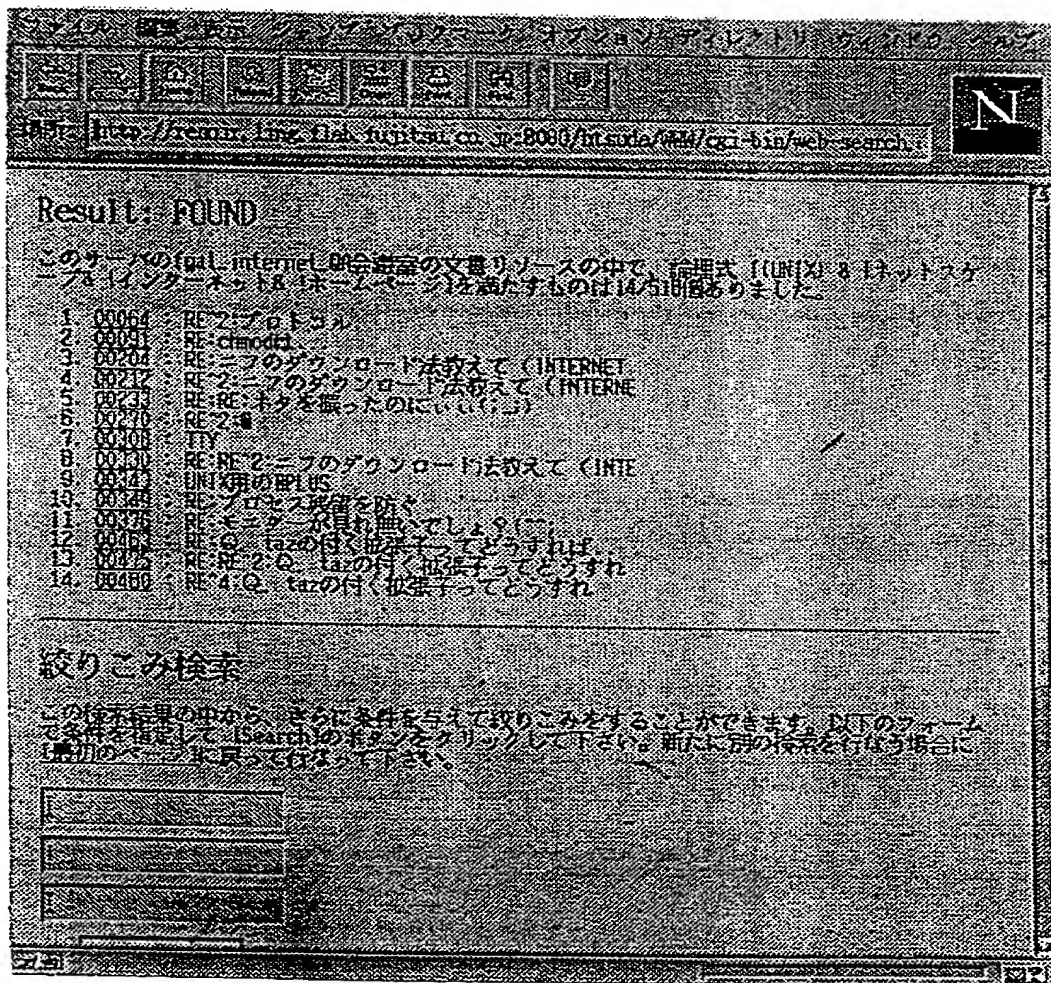
検索語:プロトコル	
会議室名: FGL インターネット	
タイトル	キーワード
A 先生がケーブルを	B 大学 C 校舎 インターネットのはじまり
プロトコル	telnet ftp UNIXマシンのプロトコル
.....
Internetの規格	プロトコルの種類 http ftp
.....

赤

【図33】

会議室内の全発言の検索(結果出力)

表示例を示す図



【図34】

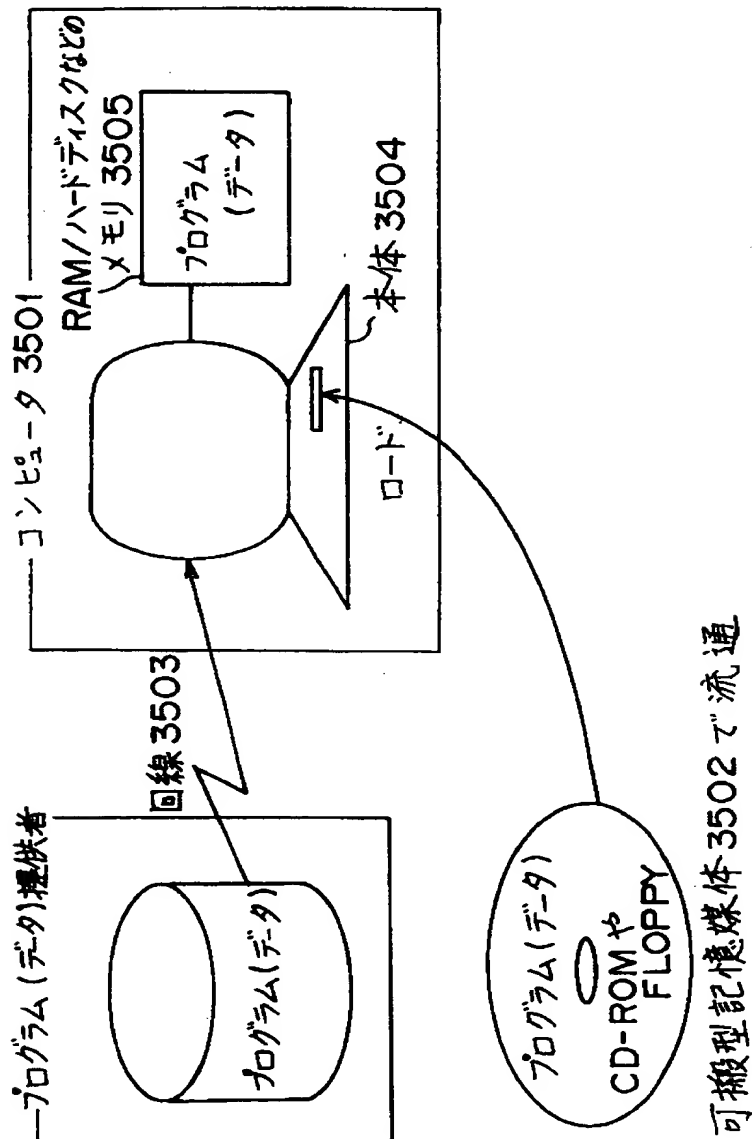
サブトピックから抽出した
キーワードの表示例を示す図

メイントピック: 何から質問すれば

サブトピック	キーワード
A 先生がケーブルを	B 大学 C 校舎 インターネットのはじまり
プロトコル	telnet ftp UNIXマシンのプロトコル
.....

【図35】

本実施の形態を実現するプログラムが
記録された記録媒体の説明図



【書類名】 要約書

【要約】

【課題】 大量の文書が含まれる文書集合からの、必要な文書又は必要な文書が含まれているであろう文書群の検索を、支援することにある。

【解決手段】 フォーラム／会議室内の文書群は、所定のサーバコンピュータ上の文書群データベース301として、蓄積される。文書群解析装置302は、文書群データベース301内の各会議室に対応する文書集合毎に、それに含まれる文書群の解析を行う。集計装置303は、文書群解析装置302による解析結果に基づいて、メタインデックス304、スレッドインデックス305、及び索引ファイル404を生成する。表示装置306は、メタインデックス304とスレッドインデックス305を用いて、キーワードビュー、スレッドビュー、又は発言者ビューの何れかの表示形態で、文書群を表示する。

【選択図】 図3

【書類名】

職権訂正データ

【訂正書類】

特許願

<認定情報・付加情報>

【特許出願人】

【識別番号】

000005223

【住所又は居所】

神奈川県川崎市中原区上小田中4丁目1番1号

【氏名又は名称】

富士通株式会社

【代理人】

申請人

【識別番号】

100074099

【住所又は居所】

東京都千代田区二番町8番地20 二番町ビル3F

大菅内外国特許事務所

【氏名又は名称】

大菅 義之

【選任した代理人】

【識別番号】

100067987

【住所又は居所】

神奈川県横浜市港北区太尾町1418-305 (

大倉山二番館) 久木元特許事務所

【氏名又は名称】

久木元 彰

特平 9-242247

出 願 人 履 歴 情 報

識別番号

[000005223]

1. 変更年月日 1996年 3月26日

[変更理由] 住所変更

住 所 神奈川県川崎市中原区上小田中4丁目1番1号

氏 名 富士通株式会社